
TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky a mezioborových studií

Studijní program: B2612 – Elektrotechnika a informatika

Studijní obor: 1802R022 – Informatika a logistika

**Odhad funkce intenzity
procesu vláken – simulační studie**

**Estimation of Intensity Function
of a Fiber Process – Simulation Study**

Bakalářská práce

Autor:	Petr Schilk
Vedoucí práce:	Mgr. Čeněk Jirsák
Konzultant:	Ing. František Kratochvíl

V Liberci 17. 5. 2013

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Petr Schilk**
Osobní číslo: **M08000302**
Studijní program: **B2612 Elektrotechnika a informatika**
Studijní obor: **Informatika a logistika**
Název tématu: **Odhad funkce intenzity procesu vláken - simulační studie**
Zadávací katedra: **Ústav řízení systémů a spolehlivosti**

Z á s a d y p r o v y p r a c o v á n í :

1. Seznamte se s matematickým pozadím úlohy (bodové procesy, proces vláken).
2. Vytvořte simulačního prostředí.
3. Porovnejte vliv parametrů na odchylku mezi teoretickou a odhadnutou funkcí intenzity.
4. Navrhněte ideální parametry pro odhad funkce intenzity.
5. Diskutujte dosažené výsledky.


Rozsah grafických prací: dle potřeby
Rozsah pracovní zprávy: cca 40 stran
Forma zpracování bakalářské práce: tištěná/elektronická
Seznam odborné literatury:

1. Beneš V., Rataj J.: Stochastic Geomerty: Selected Topics, Kluwer Academic Publishers, Boston, 2004.
2. Diggle P.J.: A kernel method for smoothing point process data. Applied Statistics (Journal of the Royal Statistical Society, Series C) 34 (1985) 138-147.
3. Dupač V., Hušková M.: Pravděpodobnost a matematická statistika. Karolinum, Praha, 1999.
4. Jirsák, Č.: Nestacionární procesy částic, diplomová práce, MFF UK, 2010.

Vedoucí bakalářské práce: **Mgr. Čeněk Jirsák**
Ústav řízení systémů a spolehlivosti
Konzultant bakalářské práce: **Ing. František Kratochvíl**
Ústav řízení systémů a spolehlivosti
Datum zadání bakalářské práce: **15. října 2012**
Termín odevzdání bakalářské práce: **17. května 2013**


prof. Ing. Václav Kopecký, CSc.
děkan




doc. Ing. Libor Tůma, CSc.
vedoucí ústavu

V Liberci dne 15. října 2012

Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Datum

17. 5. 2013

Podpis

A handwritten signature in black ink, appearing to read 'Schilch', with a stylized, flowing script.

Poděkování

Předně děkuji vedoucímu bakalářské práce Mgr. Čěnkovi Jirsákovi za účinnou metodickou, pedagogickou a odbornou pomoc i cenné rady při zpracování mé bakalářské diplomové práce. Dále zde děkuji Mgr. Manuele Mičulkové za dlouhodobou podporu a značně nekompromisní konvergenčně motivační program vedoucí ke zdárnému dokončení a odevzdání práce v řádném termínu. V neposlední řadě pak děkuji matce, otci a sestře za trpělivost v průběhu celých mých vysokoškolských studií.

Abstrakt

Bakalářská diplomová práce se zabývá numerickou aproximací funkce délkové intenzity vlákeného procesu neparametrickou metodou jádrového odhadu s následným proložením rovinou minimálního kontrastu. V rámci vlastního řešení definujeme a naprogramujeme odpovídající simulační model procesu vláken, stanovíme jeho teoretickou funkci intenzity a určíme způsoby komparace s rovinou minimálního kontrastu. Následně proběhne dostatečné množství simulačních experimentů při postupných změnách jednotlivých vstupních parametrů, statistické zpracování výsledků a porovnání absolutních/relativních chyb. V závěru práce shrneme vliv konfigurace generování vzorků vláken a doporučíme konkrétní nastavení pro vlastní jádrový odhad. Cílem studie je dosáhnout co možná nejpresnějšího odhadu funkce délkové intenzity. Praktické využití si lze představit například při zjišťování nehomogenity nanovláken.

Klíčová slova

Stochastická geometrie, proces vláken, funkce intenzity, jádrový odhad, metoda minimálního kontrastu

Abstract

The bachelor thesis deals with numerical approximation of length intensity function of a fiber process. It is using nonparametric method of kernel density estimation followed by plane interleaving with minimum contrast. Within its solutions we define and program appropriate simulation model of fiber process, we also determine the theoretical function of intensity and we identify the ways of comparison with plane of minimal contrast. Then there are thousands of simulation experiments at changes of input parameters, statistical processing of results and moreover comparison of the absolute/relative errors. In conclusion of the bachelor thesis we summarize the impact of configuration generating samples of fibers and recommend specific settings for kernel density estimation. The aim of this thesis is to achieve the most accurate approximation which is possible for function of length intensity. We can imagine its practical use for example in determining the inhomogeneous nature of the nanofibers.

Keywords

Stochastic geometry, fiber process, intensity function, kernel density estimation, minimal contrast method

OBSAH

Seznam obrázků	8
Seznam grafů.....	9
Seznam tabulek	9
1 Úvod	10
2 Náhodné procesy.....	12
2.1 Bodový proces	12
2.2 Vlastnosti bodových procesů	13
2.3 Proces vláken a jeho vlastnosti	14
2.4 Jádrový odhad funkce intenzity	15
2.4.1 Indikátor okolí bodu	17
2.4.2 Lineární převážení.....	18
2.4.3 Epanechnikovovo jádro.....	19
2.4.4 Normální rozdělení.....	20
2.5 Okrajové efekty	21
2.6 Metoda minimálního kontrastu	22
3 Simulační model.....	23
3.1 Definice modelu.....	23
3.2 Vstupní parametry modelu	24
3.3 Teoretická funkce intenzity	25
3.4 Přesnost odhadu funkce intenzity	27
4 Simulační experimenty.....	28
4.1 Střední počet vláken	28
4.2 Maximální výchylka segmentu	31
4.3 Délka segmentu	34
4.4 Rozlišení segmentu	37
4.5 Rozlišení jádrového odhadu	40
4.6 Jádrová funkce	43
4.7 Šířka pásma.....	46
4.8 Exponent minimalizace	49
4.9 Eliminace okrajových efektů	51

5	Závěry a doporučení.....	55
	Použitá literatura	57
	Přílohy	58
A	Ukázka simulačního prostředí.....	58
A.1	Parametry generovaného vzorku.....	58
A.2	Parametry jádrového odhadu	58
A.3	Generování vzorku procesu vláken	59
A.4	Aplikace jádrového odhadu	60
B	Obsah přiloženého DVD	61

Seznam obrázků

Obr. 1.1: Nanovlákná (PAN) – zvětšeno 5 000×	10
Obr. 1.2: Vzorek náhodné konfigurace vláken	11
Obr. 2.1: Dvě různé realizace rovinného bodového procesu	13
Obr. 2.2: Dvě různé realizace rovinného procesu vláken	14
Obr. 2.3: Ukázka aplikace metody jádrového odhadu	17
Obr. 2.4: Indikátor okolí bodu	18
Obr. 2.5: Lineární převážení	19
Obr. 2.6: Epanechnikovovo jádro	20
Obr. 2.7: Gaussovo jádro	21
Obr. 2.8: Ukázka vlivu okrajového efektu	21
Obr. 3.1: Cílový průběh funkce délkové intenzity	26
Obr. 4.1: Faktor středního počtu vláken	30
Obr. 4.2: Faktor maximální výchylky segmentu	33
Obr. 4.3: Faktor délky segmentu	36
Obr. 4.4: Faktor rozlišení segmentu	39
Obr. 4.5: Faktor rozlišení jádrového odhadu	42
Obr. 4.6: Faktor druhu jádrové funkce	45
Obr. 4.7: Faktor šířky pásma	48
Obr. 4.8: Ukázka analyzované konfigurace	52
Obr. 4.9: Faktor okrajových efektů (UNIFORM)	53
Obr. 4.10: Faktor okrajových efektů (TRIANGULAR)	53
Obr. 4.11: Faktor okrajových efektů (EPANECHNIKOV)	54
Obr. 4.12: Faktor okrajových efektů (GAUSSIAN)	54

Seznam grafů

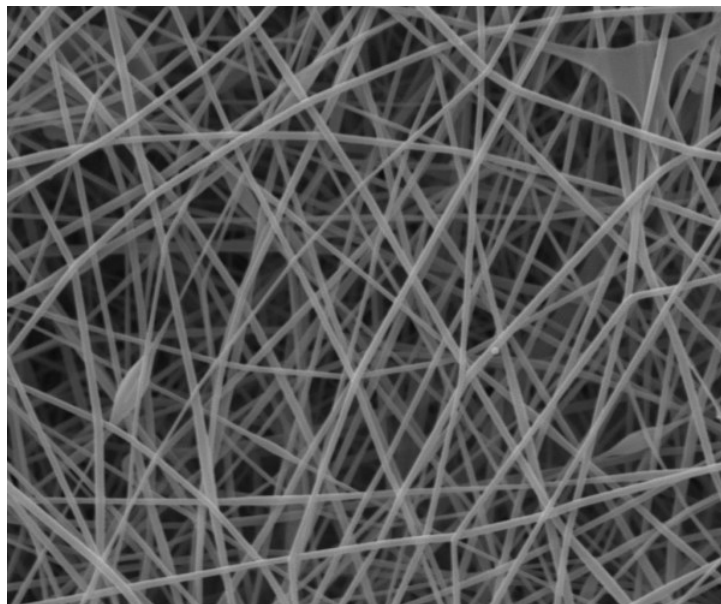
Graf 4.1: Faktor středního počtu vláken	29
Graf 4.2: Faktor maximální výchylky segmentu	32
Graf 4.3: Faktor délky segmentu	35
Graf 4.4: Faktor rozlišení segmentu	38
Graf 4.5: Faktor rozlišení jádrového odhadu	41
Graf 4.6: Faktor druhu jádrové funkce	44
Graf 4.7: Faktor šířky pásma	47
Graf 4.8: Faktor exponentu minimalizace	50
Graf 4.9: Faktor okrajových efektů.....	52

Seznam tabulek

Tab. 4.1: Faktor středního počtu vláken	29
Tab. 4.2: Faktor maximální výchylky segmentu	31
Tab. 4.3: Faktor délky segmentu	34
Tab. 4.4: Faktor rozlišení segmentu.....	38
Tab. 4.5: Faktor rozlišení jádrového odhadu	40
Tab. 4.6: Faktor druhu jádrové funkce	43
Tab. 4.7: Faktor šířky pásma	46
Tab. 4.8: Faktor exponentu minimalizace	49
Tab. 4.9: Faktor okrajových efektů.....	51

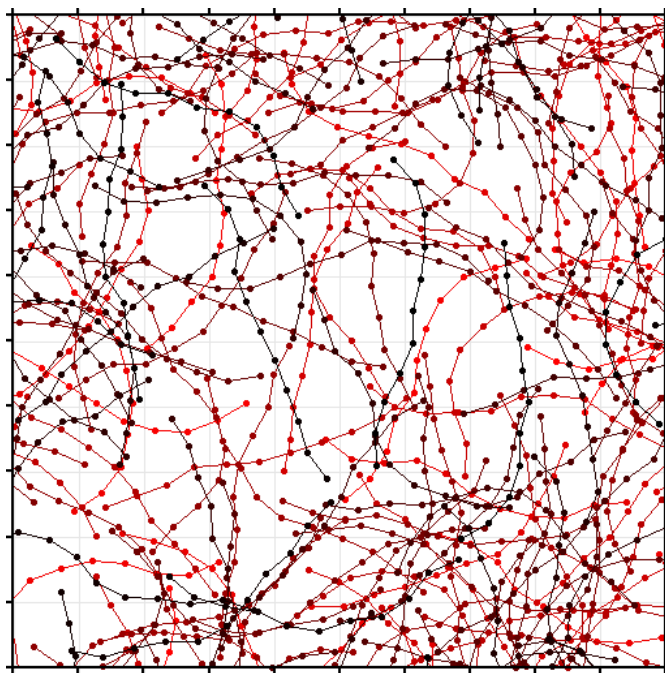
1 Úvod

Netkanými textiliemi označujeme ploché nehomogenní struktury vyrobené z náhodně orientovaných vláken. Pohybuje-li se jejich průměr v rozsahu desítek až stovek nanometrů, pak již hovoříme o nanovlákněch. Právě díky své struktuře mají nanovláknenné materiály unikátní vlastnosti a nabízí netušené možnosti pro jejich využití v mnoha oblastech – ve zdravotnictví, elektronice, strojírenství, stavebnictví; dále v chemickém, optickém, automobilovém, elektrotechnickém, kosmickém, vojenském a samozřejmě rovněž textilním průmyslu. Prostřednictvím obrázku 1.1 alespoň virtuálně pohlédneme do „objektivu“ elektronového mikroskopu.



**Obr. 1.1: Nanovlákná (PAN) – zvětšeno 5 000×
Zdroj: ELMARCO [6]**

Vlastnosti nanovláken je možné určovat různými metodami – například videometrií; tj. vyhodnocováním obrazu měřením. V rámci předkládané simulační studie se však zaměříme na partii vícerozměrné statistiky (stochastické geometrie); konkrétně nehomogenních procesů vláken na dvourozměrném euklidovském prostoru. Celou problematiku si lze jednoduše představit jako měření výskytu určitého jevu na pevně stanovené ploše normalizovaného rozměru, kterou nazýváme okno pozorování. Vizualizace konkrétní konfigurace viz obrázek 1.2. Funkce intenzity nám pak intuitivně podává obraz pravděpodobného rozložení „hmoty“ (délky) vláken v dané oblasti.



Obr. 1.2: Vzorek náhodné konfigurace vláken

Funkce intenzity může být obecně velmi složitá, těžko popsitelná a obtížně uchopitelná struktura. Jedním z tradičních způsobů, jak alespoň uspokojivě určit její průběh, je aplikace jádrového odhadu, respektive jeho numerické aproximace, a následným uplatněním metody minimálního kontrastu. Bohužel však současně nedržíme v rukách žádné exaktní vodítko, které by nám pomohlo s výběrem druhu jádra, nastavením šířky pásma, či poradilo, jak „dostatečně jemně“ generovat vlastní jádrový odhad. Ano, přesně zde se tedy otevírá prostor pro simulační studii, která se pokusí vliv jednotlivých parametrů prošetřit.

Bezprostředně následující oddíl prochází potřebné teoretické pozadí prostorové statistiky, zejména oblast (mateřských) bodových procesů, proces vláken, jádrový odhad funkce intenzity a metodu minimálního kontrastu. V dalších kapitolách vytvoříme adekvátní simulační model procesu vláken, určíme jeho teoretickou funkci intenzity a způsoby komparace s rovinou minimálního kontrastu. Následně proběhnou simulační experimenty při postupných změnách prvků množiny vstupních parametrů a statistické zpracování absolutních/relativních chyb. V závěru textu shrneme vliv konfigurace generování vzorků procesu vláken a doporučíme konkrétní nastavení pro vlastní jádrový odhad. Cílem bakalářské práce je dosáhnout co možná nejpřesnějšího odhadu funkce délkové intenzity metodou minimálního kontrastu.

2 Náhodné procesy

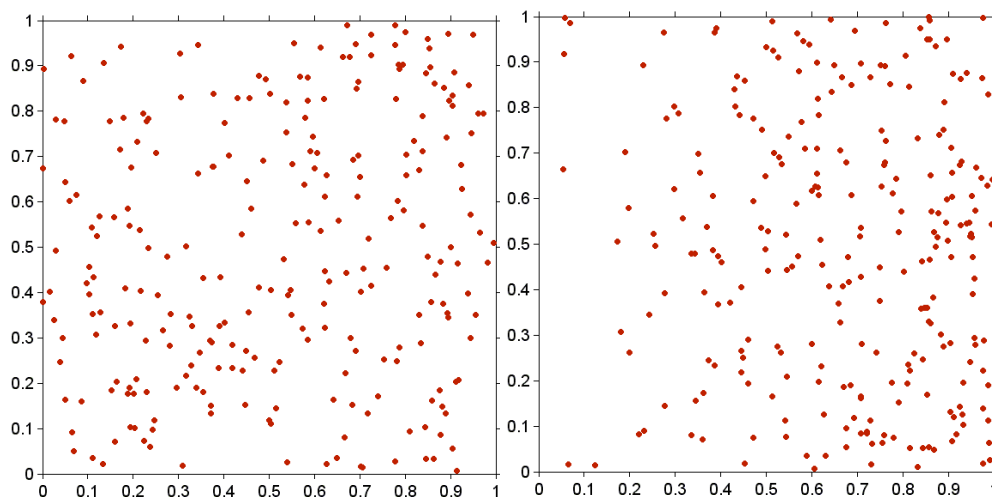
Teoretický oddíl věnujeme elementárním oblastem prostorové statistiky; konkrétně mateřskému procesu bodů a z něj vzniknuvších vláken. Dále se podíváme na neparametrickou techniku jádrového odhadu funkce intenzity s následnou aplikací metody minimálního kontrastu.

2.1 Bodový proces

Je to právě bodový proces (*point process*), jehož realizace nám vytváří stabilní podhoubí náhodné konfigurace vláken. Pro účely tohoto materiálu, budování simulačního modelu a v rámci realizovatelnosti názorných experimentů se omezíme na procesy v dvourozměrném eukleidovském prostoru; prakticky tedy pracujeme výhradně s rovinnými (pseudo)náhodnými daty. Pro analogické jevy a úplné matematické pozadí v prostorech vyšších dimenzí viz [8].

Bodový proces φ lze obecně definovat jako měřitelné zobrazení, jak podrobněji rozepisuje například Pawlas [9], avšak pohodlně vystačíme i se zjednodušenou představou, že jde o stochastickou bodovou konfiguraci – náhodné rozmístění izolovaných bodů v rovině, často pevně ohraničené. Označíme jej jako konečný, jestliže obsahuje konečně mnoho bodů. Střední počet takových bodů v rámci konkrétní množiny A označujeme symbolem $\varphi(A)$. Můžeme si též dovolit předpokládat realizaci pouze tzv. jednoduchých bodových procesů, kdy instance φ obsahují pouze navzájem různé body. Nepřipouštíme tímto výskyt bodů s větší násobností.

V případě dimenze vyšší než jedna, nemá označení „proces“ žádnou souvislost s dynamickým vývojem v čase. Zajímá nás jen a pouze rozmístění bodů v dvourozměrném eukleidovském prostoru, a to v pevně stanoveném časovém okamžiku. Proto se jakožto vhodnější označení nabízí spíše slovní spojení „bodové pole“. Na obrázku 2.1 jsou zobrazeny dvě konkrétní realizace rovinných bodových procesů s kardinalitou 256 ve čtvercovém okně standardního (tj. normovaného) rozměru $[0, 1]^2$ – délka strany čtverce odpovídá 1 j (jedné jednotce). Identické okno pozorování (*observation window*) poslouží ve funkci „kukátka“ do světa stochastické geometrie i v případě simulačního modelu, který v pozdější fázi definujeme a naprogramujeme; umožní nám vizuální porovnání výsledků jednotlivých experimentů.



Obr. 2.1: Dvě různé realizace rovinného bodového procesu

2.2 Vlastnosti bodových procesů

Nejvýznamnější charakteristikou bodových procesů je jejich pravděpodobnostní rozdělení/rozložení Π , které vyjadřuje způsob náhodného rozmístění (tj. distribuci) bodů v rámci dané plochy. Již na obrázku 2.1 jsme si mohli všimnout konkrétních vzorků s rozložením rovnoměrným a trojúhelníkovým. Bodový proces φ a jeho pravděpodobnostní rozdělení Π se často ztotožňují.

Pokud je příslušné pravděpodobnostní rozdělení invariantní vůči posunutí, jedná se o stacionární bodový proces. Jestliže je neměnné vůči rotaci kolem počátku souřadné soustavy, hovoříme o procesu izotropním.

Míra intenzity (*intensity measure*) $\Lambda(A)$ bodového procesu φ vyjadřuje střední počet bodů na dané množině A :

$$\Lambda(A) = \mathbb{E}\varphi(A). \quad (2.1)$$

Existuje-li hustota λ míry Λ , potom λ se nazývá funkce intenzity (*intensity function*); formálně tuto skutečnost zapíšeme ve formě integrace funkce intenzity přes celou množinu A takto:

$$\Lambda(A) = \int_A \lambda(x) dx. \quad (2.2)$$

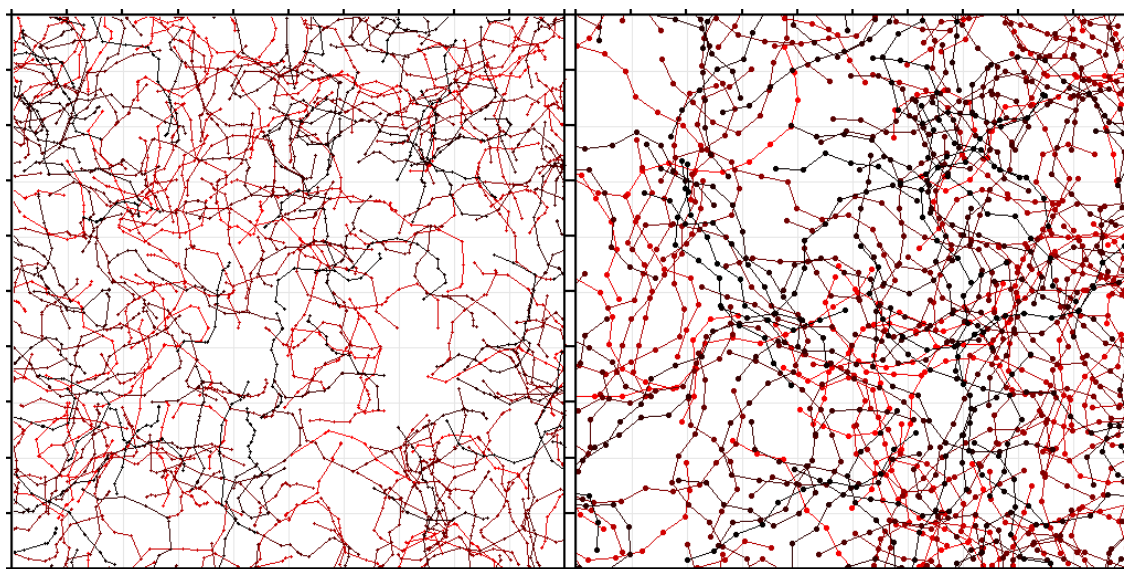
Jak píše Pawlas [9], podstatu funkce intenzity můžeme intuitivně chápat tak, že $\lambda(x) dx$ je pravděpodobnost výskytu bodu procesu v infinitezimálně „malém“ okolí x

s objemem dx . Jestliže funkce intenzity λ existuje a tato je na dané množině konstantní, pak příslušný bodový proces označíme jako homogenní (*homogenous*).

2.3 Proces vláken a jeho vlastnosti

Generalizací bodového pole přejdeme k procesu vláken (*fiber process*¹). Procesem vláken Φ tedy analogicky rozumíme jejich stochastickou konfiguraci – náhodné rozmístění vláknenných útvarů v rovině, často účelně pevně vymezené oknem pozorování W . V případě bodového procesu jsme se zaměřili na (střední) počty bodů, avšak u vláken nás zajímá především jejich celková délka v dané oblasti. Střední délku vláken v rámci konkrétní množiny W označíme symbolem $\Phi(W)$. Opět lze předpokládat realizaci pouze tzv. jednoduchých vláknenných procesů, kdy konkrétní vzorky Φ obsahují pouze navzájem různá vlákna; nepřipouštíme jejich násobnost.

Na obrázku 2.2 si ve standardních oknech pozorování $[0, 1]^2$ prohlédneme ukázky dvou různých realizací rovinných vláknenných procesů, obě se střední délkou 64. Jednotlivé instance vláken mohou být jak náhodné délky (vlevo), tak mít délku pevně danou (vpravo jsou vesměs všechna vlákna dlouhá 0,5 j; jednotky viz pododdíl 2.1). Ohledně zbarvení – čím tmavší má vlákno barvu, tím dříve bylo algoritmem vygenerováno. Jasně červená vláknina je takříkajíc „nejmladší“, vyrostla až jako poslední.



Obr. 2.2: Dvě různé realizace rovinného procesu vláken

¹ Není bez zajímavosti, že v anglicky psané literatuře se běžně objevují spojení „fiber process“ i „fibre process“. Slůvko „fibre“ bychom však ve slovníku americké angličtiny hledali marně – Britové jej převzali pravděpodobně z francouzštiny.

I v případě procesu vláken pracujeme s pravděpodobnostním rozdělením/rozložením Π , které se zde vztahuje k mateřskému procesu bodů (*parent point process*). Mateřské body plní funkci zárodků vláken, de facto hovoříme o jejich počátečních bodech. Samotné tělo vlákna pak vyrůstá z mateřského bodu a jeho distribuce je rovnoměrná – v kruhovém okolí (zárodku) s poloměrem o velikosti maximální možné délky vlákna. Na obrázku 2.2 si vlevo povšimněme vzorku s rozložením rovnoměrným, vpravo pak trojúhelníkovým. Proces vláken Φ a jeho pravděpodobnostní rozdělení Π se často ztotožňují. Podobně jako u bodového pole, i proces vláken může být ve speciálních konstelacích stacionární, případně izotropní povahy.

Míra délkové intenzity (*length intensity measure*) $\Lambda(W)$ vlákenného procesu Φ vyjadřuje střední délku vláken na dané ploše W ; přehledně přepíšeme:

$$\Lambda(W) = \mathbb{E}\Phi(W). \quad (2.3)$$

Analogicky k předchozímu – existuje-li hustota λ míry Λ (a v našem případě existovat bude), potom λ se nazývá funkce délkové intenzity (*length intensity function*); formálně zapíšeme ve formě dvojné integrace λ přes okno pozorování W :

$$\Lambda(W) = \iint_W \lambda(x, y) dx dy. \quad (2.4)$$

Podstatu funkce délkové intenzity tedy zjednodušeně chápeme tak, že $\lambda(x, y) dx dy$ je průměrná (střední) délka všech vláken v infinitezimálně „malém“ okolí bodu $[x, y]$ s objemem $dx dy$. Jestliže je funkce délkové intenzity λ na dané ploše konstantní, pak příslušný proces vláken označíme jako homogenní. Takové procesy však nejsou příliš „zajímavé“, a proto bude jistě vhodnější postavit simulační model na procesu nehomogenním (*inhomogeneous*).

2.4 Jádrový odhad funkce intenzity

Konvenční vyšetření průběhu funkce délkové intenzity může být zejména u nehomogenních procesů vláken víceméně neřešitelnou úlohou. Praktická potřeba statistického zpracování vzorků nás tímto vede k užití neparametrických metod, respektive softwarové implementaci jejich numerických aproximací (za předpokladu akceptace jistých chyb/odchylek). Zřejmě nejčastěji se setkáváme s aplikací metody jádrového odhadu (*KDE, kernel density estimation*).

Princip metody jádrového odhadu funkce délkové intenzity si lze zjednodušeně představit tak, že „hmotu“ vláken (nekonečně velké výšky / nekonečně malé šířky) rozprostřeme podle pravděpodobnostní hustoty zvoleného druhu jádra do jejich okolí, jehož velikost určuje šířka pásma (*bandwidth*). Postupně tedy integrujeme jádrovou funkcí převážené body okna pozorování, které jsou „blízko“ středu jádra; čím blíže vláknu se nacházíme, tím vyšší váhu obdržíme. Matematicky bychom výše uvedenou představu mohli vyjádřit následujícím vzorcem:

$$\hat{\lambda}(x, y) = \Delta l \cdot \iint_w h_b \sqrt{(x - x_h)^2 + (y - y_h)^2} dx dy, \quad (2.5)$$

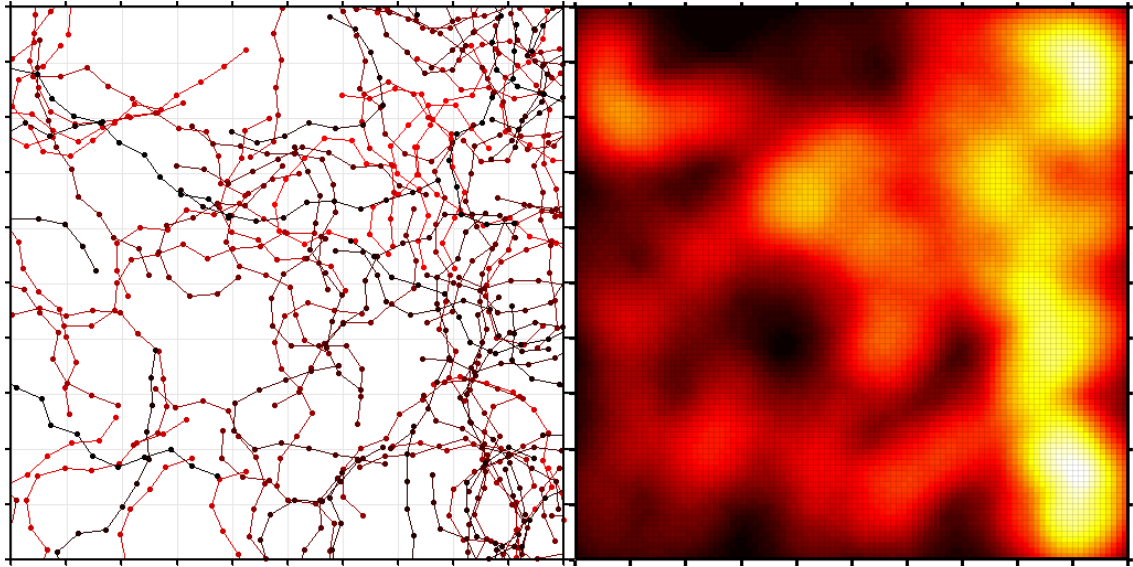
kde $\hat{\lambda}(x, y)$ označuje odhad funkce délkové intenzity vlákenného procesu, h_b je předem navolená jádrová funkce (*kernel function*) s šířkou pásma b a $\sqrt{(x - x_h)^2 + (y - y_h)^2}$ vyjadřuje eukleidovskou vzdálenost bodu $[x, y]$ okna pozorování od středu jádra $[x_h, y_h]$. Výsledek je ještě nutné převážít koeficientem Δl , protože jedno jádro běžně nereprezentuje celé vlákno, ale pouze jeho malou část, která má délku právě Δl . Raději zdůrazněme, že po celé délce vlákna rozmísťujeme jader hned několik, a to v přiměřené hustotě zejména vzhledem k šířce pásma.

Jádrovou funkci h_b vybíráme obvykle takovou, aby byla symetrická podle počátku, kladná pouze na jistém středovém okolí (určeném šířkou pásma), nabývala svého maxima v nule a byla ve všech směrech ke středu neklesající. V dalších oddílech tohoto textu zjistíme, že odhad je poměrně značně citlivý na velikost šířky pásma. Jádrové funkce bývají standardně definované pro šířku pásma rovnu jedné. Jak však uvádí Jirsák [8], lze je velmi snadno přeskálovat na libovolně široké okolí podle vztahu:

$$h_b(x) = \frac{h_1(x/b)}{b^2}, \quad (2.6)$$

kde b značí šířku pásma.

Následuje podrobnější popis charakteristik čtyř nejběžněji užívaných jader, poťazmo příslušných jádrových funkcí. Ještě předtím si však na obrázku 2.3 demonstřujeme aplikaci metody jádrového odhadu na konkrétním vzorku – vlevo vidíme analyzovanou konfiguraci vláken, vpravo je vizualizace jádrového odhadu, tzv. spektrální mapa. Čím je barva spektra světlejší, tím vyšší je v daném místě odhad délkové intenzity.



Obr. 2.3: Ukázka aplikace metody jádrového odhadu

2.4.1 Indikátor okolí bodu

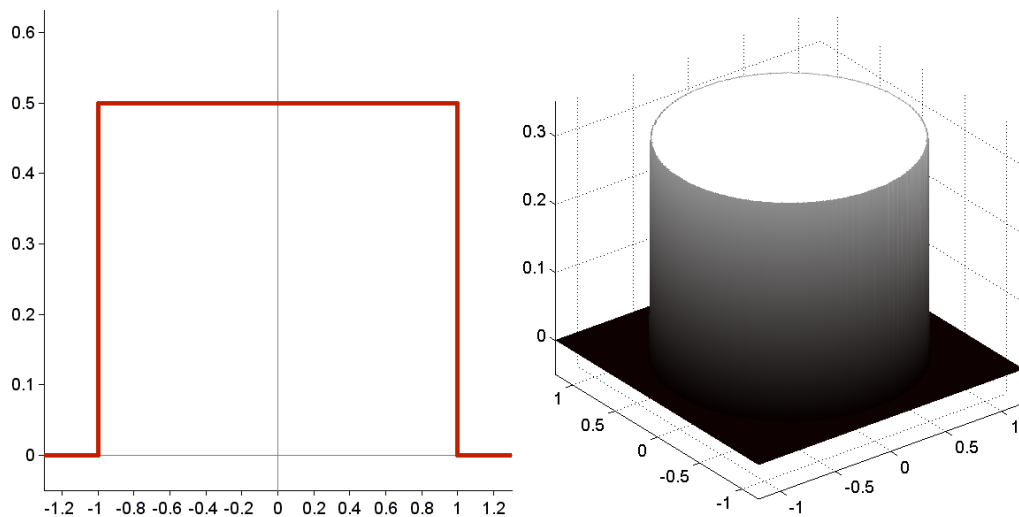
Jádrovou funkci indikátoru okolí bodu (*uniform kernel function*) tvoří hustota rozložení pravděpodobnosti rovnoměrného rozdělení. Samotné jádro nabývá podoby rotačního válce, jehož objem musí mít (z důvodu zachování podstaty pravděpodobnostního pozádí) velikost právě 1. Předpis funkce pro rovinná data s šířkou pásma jedné jednotky je následující:

$$h_1(x, y) = \frac{1}{\pi} 1_{\{\|u\| \leq 1\}}, \quad (2.7)$$

kde norma $\|u\| = \sqrt{(x - x_h)^2 + (y - y_h)^2}$ má význam eukleidovské vzdálenosti du $[x, y]$ okna pozorování od středu $[x_h, y_h]$ aplikovaného jádra. Funkce $1_{\{\|u\| \leq 1\}}$ je indikátor množiny – je-li podmínka splněna (patří-li prvek do množiny), pak vrátí hodnotu jedna, v opačném případě nula. Provedeme-li přeškálování dle vzorce (2.6), dostaneme výhodný tvar pro libovolně volenou šířku pásma b :

$$h_b(x, y) = \frac{1}{\pi b^2} 1_{\{\|u\| \leq b\}}. \quad (2.8)$$

Velikost a tvar jader s šířkou pásma rovnou jedné jednotce pro jedno a dvourozměrná data vidíme na obrázku 2.4. Indikátor okolí bodu patří mezi vůbec nejčastěji používané jádrové funkce při řešení problematiky odhadů intenzity náhodných procesů, zejména pro svoji jednoduchost, potažmo minimální náročnost na výpočetní výkon.



Obr. 2.4: Indikátor okolí bodu

2.4.2 Lineární převážení

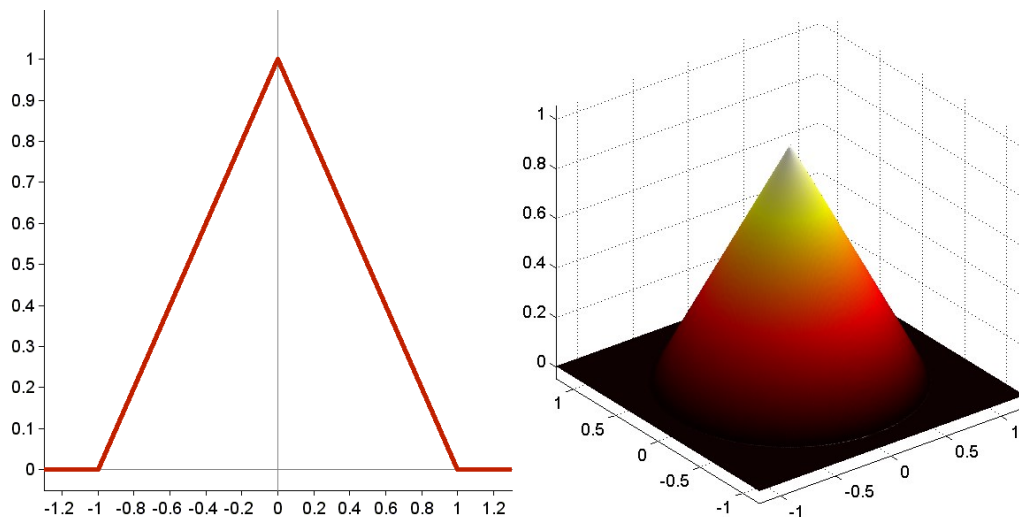
Základem lineárního převážení (*triangular kernel function*) je hustota rozložení pravděpodobnosti rozdělení trojúhelníkového. Bližší body z rozsahu šířky pásma obdrží adekvátně (lineárně) vyšší váhy, než body vzdálenější. Vlastní jádro dvourozměrných dat tvoří rotační kužel, který analogicky splňuje podmínku velikosti objemu rovnu jedné:

$$h_1(x, y) = \frac{3}{\pi} (1 - \|u\|) 1_{\{\|u\| \leq 1\}}. \quad (2.9)$$

Norma $\|u\| = \sqrt{(x - x_h)^2 + (y - y_h)^2}$ má význam eukleidovské vzdálenosti du $[x, y]$ okna pozorování od středu $[x_h, y_h]$ aplikovaného jádra. Funkce $1_{\{\|u\| \leq 1\}}$ záměrně nuluje výslednou hodnotu pro body mimo dosah šířky pásma. Provedeme-li přeškálování dle vzorce (2.6), dostaneme tvar pro libovolně volenou šířku pásma b :

$$h_b(x, y) = \frac{3}{\pi b^2} \left(1 - \frac{\|u\|}{b}\right) 1_{\{\|u\| \leq b\}}. \quad (2.10)$$

Velikost a tvar jader s šířkou pásma rovnu jedné jednotce pro jedno a dvourozměrná data názorně vykresluje obrázek 2.5. Lineární převážení je též hojně užívaným typem jádra a nesmí tak chybět ani v námi vytvořeném simulačním prostředí. Pouhou vizuální analýzou spektrální mapy jej lze poměrně snadno odlišit od indikátoru okolí bodu, protože výsledný odhad funkce intenzity nevytváří tak ostré (kontrastní) přechody a výrazné hrany.



Obr. 2.5: Lineární převážení

2.4.3 Epanechnikovovo jádro

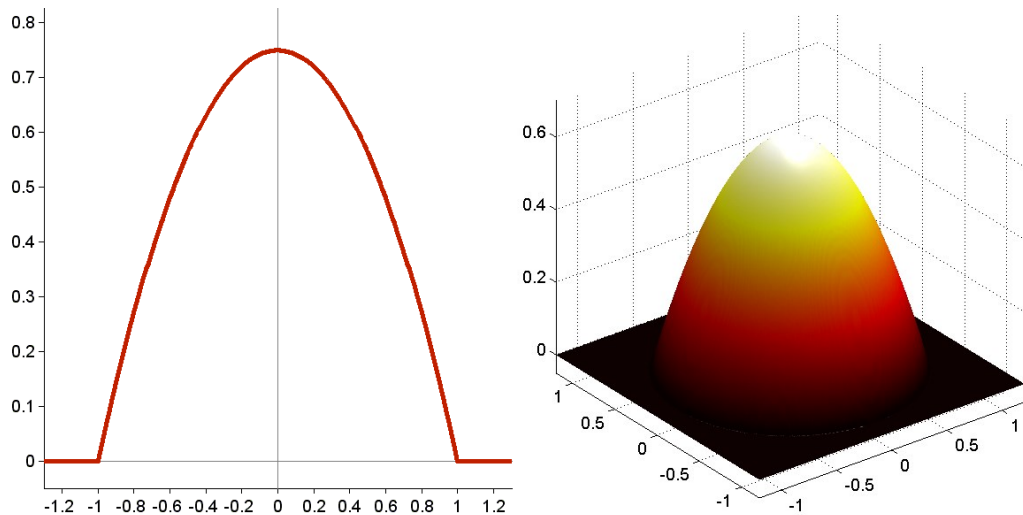
Epanechnikovovo jádro (*Epanechnikov kernel function*) spatřilo poprvé světlo světa v lednu roku 1969, kdy jej prostřednictvím příspěvku ve vědeckém časopisu publikoval moskevský statistik V. A. Epanechnikov [7]. Jedná se o kvadratické převážení – na dvourozměrná data aplikujeme jádro ve tvaru části rotačního paraboloidu, pro nějž platí obdobná podmínka ohledně velikosti objemu rovnu jedné:

$$h_1(x, y) = \frac{2}{\pi} (1 - \|u\|^2) 1_{\{\|u\| \leq 1\}}. \quad (2.11)$$

Norma $\|u\| = \sqrt{(x - x_h)^2 + (y - y_h)^2}$ má opět význam eukleidovské vzdálenosti bodu $[x, y]$ okna pozorování od středu $[x_h, y_h]$ aplikovaného jádra. Funkce $1_{\{\|u\| \leq 1\}}$ nuluje výslednou hodnotu pro body mimo dosah šířky pásma. Provedeme-li přeškálování dle vzorce (2.6), dostaneme tento tvar pro libovolně volenou šířku pásma b :

$$h_b(x, y) = \frac{2}{\pi b^2} \left[1 - \left(\frac{\|u\|}{b} \right)^2 \right] 1_{\{\|u\| \leq b\}}. \quad (2.12)$$

Situace na přímce a v rovině je vykreslena na obrázku 2.6. Jádro má strmý průběh a ostrý přechod na rozhraní okraje šířky pásma. Vizuálním porovnáváním spektrálních map jej však prakticky nejsme schopni rozeznat od lineárního převážení.



Obr. 2.6: Epanechnikovovo jádro

2.4.4 Normální rozdělení

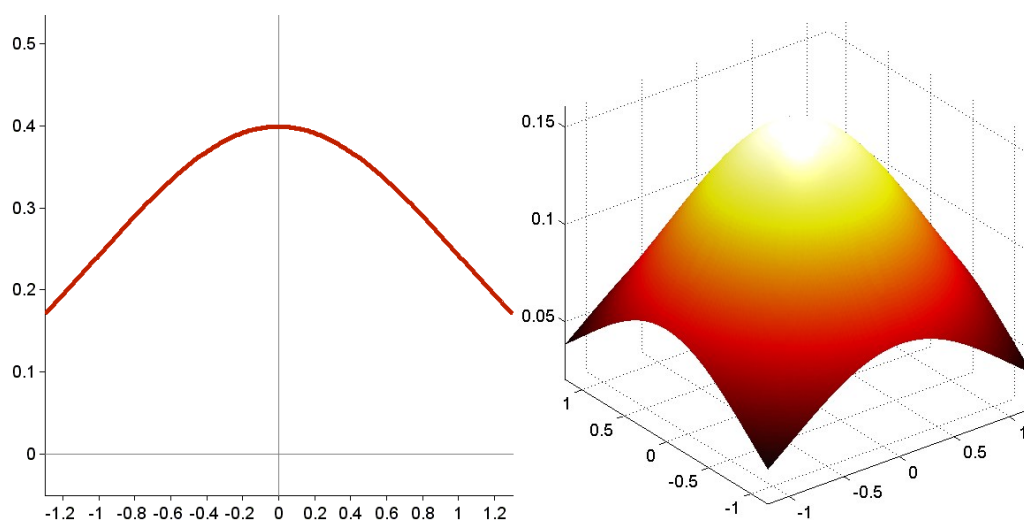
Jakožto jádro lze použít dokonce i takovou funkci, která na celém svém definičním oboru nabývá vesměs kladných hodnot (má tedy šířku pásma de facto nekonečnou) – například hustotu rozdělení pravděpodobnosti normálního/Gaussova rozdělení (*Gaussian kernel function*). Pro bivariantní normální rozložení s rozptylem jedné jednotky podél obou souřadných os platí tento všeobecně známý předpis:

$$h_1(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}\|u\|^2}. \quad (2.13)$$

Norma $\|u\| = \sqrt{(x - x_h)^2 + (y - y_h)^2}$ má význam eukleidovské vzdálenosti du $[x, y]$ okna pozorování od středu $[x_h, y_h]$ aplikovaného jádra. Provedeme-li přeškálování dle vzorce (2.6), dostaneme vzorec pro libovolně volený rozptyl b :

$$h_b(x, y) = \frac{1}{2\pi b^2} e^{-\frac{\|u\|^2}{2b^2}}. \quad (2.14)$$

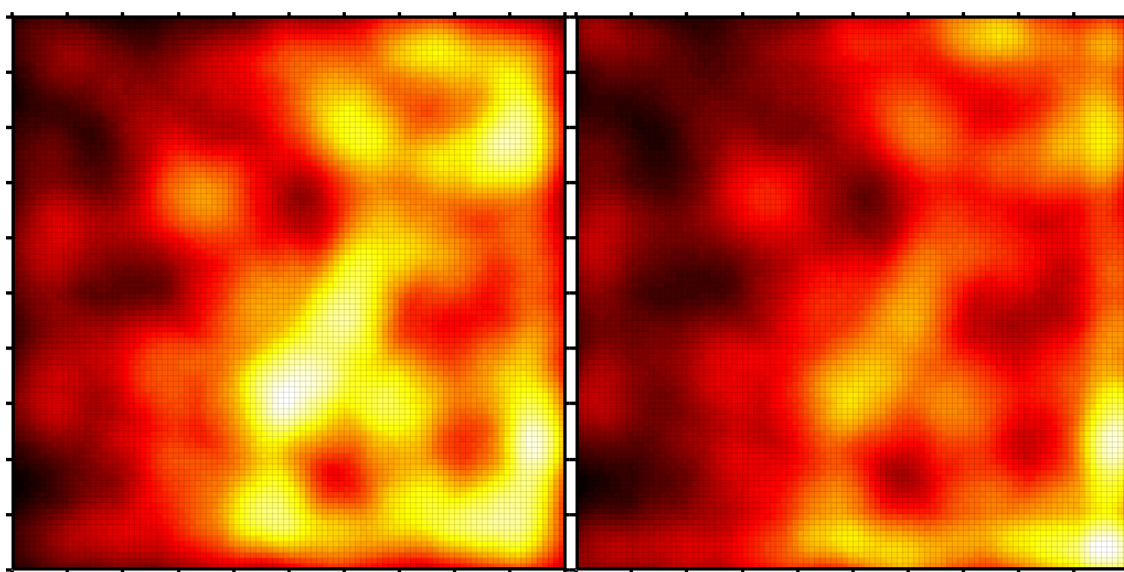
Jak můžeme odtušit již z ilustračního obrázku 2.7, při volbě velké šířky pásma (respektive rozptylu) bude nejspíš aplikace Gaussova jádra trpět příliš razantním vyhlazením spektrální mapy, neboť všechny body okna pozorování obdrží kladný přírůstek odhadu od každého jádra; bez ohledu na přímou vzdálenost. Přesto bude jistě velmi zajímavé vyhodnocení výsledků s nastaveným úzkým pásmem a při současné aktivaci eliminace okrajových efektů – viz dále.



Obr. 2.7: Gaussovo jádro

2.5 Okrajové efekty

Surový (nekorigovaný) jádrový odhad viditelně trpí ztrátou intenzity podél hranic okna pozorování, což je jev označovaný jako okrajový efekt (*edge effect*, *boundary bias*) – viz situaci vlevo na obrázku 2.8 (zde je konkrétně aplikováno Epanechnikovovo jádro s šířkou pásma 0,1 j). Tato ztráta informace je způsobena skutečností, že ačkoli neuvažujeme (a často ani neznáme) konfiguraci vláken mimo hranice okna, tak všechna jádra (hypoteticky) umístěná právě mimo okno, by mohla (v závislosti na šířce pásma) svým objemem obsah okna pozorování ovlivnit. Působnost jevu sahá od hranic okna pozorování až do vzdálenosti šířky pásma směrem ke středu vyšetřované oblasti.



Obr. 2.8: Ukázka vlivu okrajového efektu

Korekci okrajových efektů lze v teoretických modelech řešit modelováním do (o šířku pásma) rozšířeného okna pozorování, v praktické implementaci je nutné použít koeficienty – každý bod $[m, n]$ vyšetřované oblasti bychom korigovali následující vahou C :

$$C(m, n) = \frac{1}{\iint_W h(m, n) dx dy}, \quad (2.15)$$

kde $h(m, n)$ je příslušná jádrová funkce se středem v bodě $[m, n]$. Jádra jsou symetrická a jejich objem je roven jedné jednotce. Integrací jádrové funkce v okně pozorování tedy zjistíme, jaké množství informace ztrácíme vně okna a inverzí pak vlastní koeficient. Spektrální mapu téže konfigurace vláken po eliminaci okrajových efektů (převážním bodů) vidíme v pravé části obrázku 2.8. Dodejme ještě, že při užití Gaussova jádra dochází, vzhledem k charakteru jádrové funkce, ke korekci všech bodů okna pozorování.

2.6 Metoda minimálního kontrastu

Princip regresní metody minimálního kontrastu (*minimum contrast method*) spočívá v aproximaci funkce délkové intenzity pomocí roviny, kterou prokládáme jádrový odhad takovým způsobem, aby jeho kontrast vůči této rovině byl minimální možný. Označíme-li rovinu minimálního kontrastu

$$\lambda_{mc}(x, y) = a \cdot x + b \cdot y + c, \quad (2.16)$$

pak hledáme takové reálné koeficienty a, b, c , aby výraz

$$\iint_W |\lambda_{mc}(x, y) - \hat{\lambda}(x, y)|^p dx dy, \quad (2.17)$$

nabýval minimální hodnoty. Funkce $\hat{\lambda}(x, y)$ zde představuje jádrový odhad, absolutní hodnota kontrast a exponent minimalizace p volíme z přirozených čísel – nejčastěji se setkáváme s hodnotou dvě; v takovém případě se jedná o metodu nejmenších čtverců.

3 Simulační model

V rámci třetího oddílu definujeme simulační model rovinného procesu vláken, jehož numerickou aproximaci implementujeme v softwarovém prostředí MATLAB® dodávaném společností MathWorks, Inc. Technicky bude vlastní simulace realizována algoritmem třídy Monte Carlo. V dalším textu statisticky porovnáme přesnost získaných odhadů vůči teoretickému průběhu funkce délkové intenzity, a to při postupných změnách jednotlivých parametrů.

3.1 Definice modelu

Uvažujme proces vláken v \mathbb{R}^2 ; eukleidovské rovině. Simulační okno pozorování W nechť má standardní (tj. normovaný) rozměr čtverce o velikosti $[0, 1]^2$. Jednotlivé instance vláken budeme generovat modifikovaným algoritmem náhodné procházky (*random walk*); vlákna jsou tedy suplována (interpolována) lomenými čarami.

Nejprve spustíme mateřský bodový proces zárodků. Ve směru osy x distribuujeme body (přesněji jejich x -ové souřadnice) podle hustoty trojúhelníkového rozložení pravděpodobnosti – v praktické realizaci je použito rovnoměrné rozdělení s následnou transformací zamítací metodou. Nejnižší intenzitu bychom naměřili při levé hranici okna, nejvyšší zcela vpravo a růst je lineárního charakteru. Souřadnici y generujeme rovnoměrně po celé šíři okna. Teoretická funkce intenzity se bude ve směru osy y držet konstantních hladin.

Z předpřipravených výchozích bodů (zárodků) určíme izotropně další bod (uzel); body jsou vždy spojeny úsečkou a takto vygenerovány další segmenty až do vytvoření celé (požadované) délky vlákna. Pro každý bod (uzel) je prakticky nutné určit rovnoměrně náhodně velikost úhlu, pod kterým pokračuje bezprostředně navazující segment vzhledem k předchozímu, a následně nanese požadovanou délku segmentu, čímž získáme další uzel (následníka). Vlastní „hmota“ vláken je tímto způsobem symetricky rozdělována v okolí zárodečných bodů mateřského procesu.

Abychom zachovali konzistentní hustotu vláken i v blízkosti hranic okna pozorování, je nezbytné zvážit rozšíření vymezené plochy, ve které realizujeme mateřský bodový proces, i mimo hranice okna pozorování. Vlákna vně tak mohou volně vstupovat do okna, vnitřní z něj vystupovat a případně se dokonce vracet zpět. Rozšíříme si proto okno o 0,5 j na všechny čtyři strany směrem od středu. Prakticky budeme modelovat do

ohraničené části roviny, která má čtyřikrát větší plochu než samotné okno pozorování. Zároveň tím vzniká omezení na maximální délku celého vlákna, a to právě na 0,5 j.

3.2 Vstupní parametry modelu

Alespoň stručně popíšme možnosti nastavení jednotlivých vstupních parametrů našeho modelovacího prostředí. V průběhu každé simulační úlohy postupně měníme jednu z konfiguračních položek a zkoumáme její vliv na přesnost odhadu metodou minimálního kontrastu; zbytek sady zůstává neměnný. Určení teoretické funkce intenzity a způsob komparace s rovinou odhadu bezprostředně následují v dalších pododdílech.

Střední počet vláken n_f je průměrný počet vláken, jejich mateřský bod spadá do čtverce pozorování. Potřebujeme-li však obsah okna generovat automaticky podle teoretické funkce intenzity, nastavíme hodnoty minimální a maximální intenzity λ_{min} a λ_{max} .

Konfiguraci vláken ovlivníme určením počtu segmentů (úseček) n_s , jimiž vesměs všechna vlákna interpolujeme a pevnou délkou každého takového segmentu l_s . Vzdálenost uzlů však může být v případě potřeby generována i náhodně, a to s rovnoměrným rozdělením pravděpodobnosti na intervalu $(0, l_{max})$. Maximální úhlovou výchylku (na obě strany) dvou sousedních segmentů lze nastavit pomocí amplitudy δ_s .

Jádrové funkce máme připravené celkem čtyři – konkrétně všechna již v předchozím textu charakterizovaná jádra: indikátor okolí bodu, lineární převážení, Epanechnikovo a Gaussovo jádro. Šířku pásma jádrové funkce b je možné nastavit libovolně až do velikosti 0,5 j.

Jádrový odhad funkce intenzity numericky strádáme ve čtvercové matici, jejíž prvky reprezentují mřížku okna pozorování. V podstatě jde o jemnost kroku (granularitu) numerické integrace jádrových funkcí. Minimální rozlišení výstupní matice je rozměr 10×10 (odpovídá zrnitosti 0,1 j). Rozlišením segmentu rozumíme skutečnost, s jakou jemností rozmístíme jádra na jednotlivé úsečky lomených čar vláken. Nastavíme-li hodnotu jedna, pak každý segment reprezentují dvě jádra (v počátečním a koncovém bodě úsečky). Rozlišení „2“ dělí segment na dva sektory – jádra aplikujeme celkem tři. Jádrové funkce je přirozeně nutné korigovat délkou sektoru, jehož „hmotu“ jádro reprezentuje – viz Δl ve vzorci (2.5).

Exponent minimalizace p může být jakékoli přirozené číslo, vystupuje v minimalizační sumě (2.17). Eliminaci okrajových efektů lze přepínačem aktivovat/vypnout; za-

pnutím dochází k převážení všech prvků výstupní matice jádrového odhadu váhami numericky integrovanými dle předpisu (2.15).

3.3 Teoretická funkce intenzity

Vzhledem ke způsobu, jakým je realizován proces vláken, bude mít teoretická funkce intenzity $\lambda(x, y)$ podobu nakloněné roviny – minimální intenzitu předpokládáme při levé hranici okna, nejvyšší pak zcela vpravo a růst je lineárního charakteru. Ve směru osy y se hustota drží konstantních hladin, a proto má rovnice tvar, který je invariantní vůči y -ové souřadnici:

$$\lambda(x, y) = a \cdot x + b, \text{ kde } a, b \in \mathbb{R}. \quad (3.1)$$

Reálné koeficienty a, b se pokusíme nalézt nejprve postupnou integrací cílové funkce intenzity:

$$\Lambda([0, 1]^2) = \int_0^1 \int_0^1 \lambda(x, y) dx dy, \quad (3.2)$$

$$\Lambda([0, 1]^2) = \int_0^1 \int_0^1 (a \cdot x + b) dx dy, \quad (3.3)$$

$$\Lambda([0, 1]^2) = \int_0^1 (a \cdot x + b) dx, \quad (3.4)$$

$$\Lambda([0, 1]^2) = \left[\frac{a}{2} \cdot x^2 + b \cdot x \right]_0^1 = \frac{a}{2} + b. \quad (3.5)$$

Dále využijeme vztahu $\Lambda([0, 1]^2) = n_f \cdot n_s \cdot l_s$, ve kterém n_f značí střední počet vláken (respektive mateřských bodů) v okně pozorování, n_s počet segmentů každého vlákna a l_s je pevná délka segmentů. Generujeme-li délku úseček náhodně, pak bychom místo l_s dosadili střední hodnotu, prakticky tedy $0,5 \cdot l_{max}$.

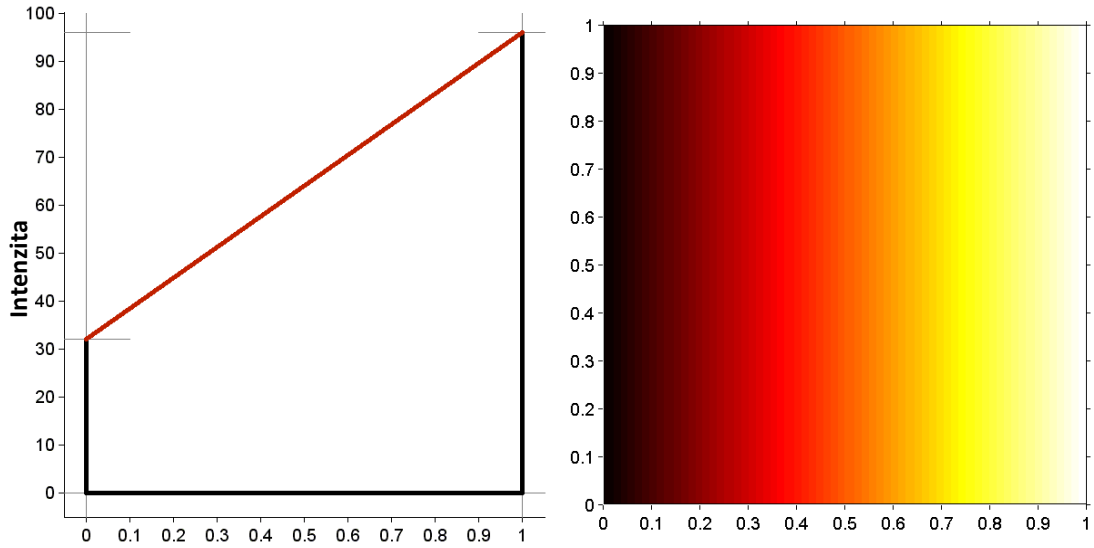
Poslední informací nutnou k dopočítání koeficientů je skutečnost, že teoretická funkce intenzity nabývá hodnoty nula (protíná rovinu okna pozorování) na přímce $x = -0,5$; a to z toho důvodu, že jsme si kvůli zachování konzistence hustoty předsadili čtverec pozorování právě o velikost 0,5 j. Dosazením do obecného tvaru (3.1) dostáváme s drobnou úpravou vztah mezi oběma koeficienty $b = \frac{a}{2}$ a vzhledem k výsledku integrace (3.5) platí $a = \Lambda([0, 1]^2)$.

Konečné podoby teoretické funkce intenzity našeho simulačního modelu vypadají následovně:

$$\lambda(x, y) = (n_f \cdot n_s \cdot l_s) \cdot x + \frac{1}{2} \cdot (n_f \cdot n_s \cdot l_s), \quad (3.6)$$

$$\lambda(x, y) = (n_f \cdot n_s \cdot l_s) \cdot \left(x + \frac{1}{2}\right). \quad (3.7)$$

Ukázkový průběh cílové funkce intenzity ilustruje obrázek 3.1; vlevo je nárys, vpravo spektrální mapa. Střední počet mateřských bodů byl 128, konfigurace vláken nastavena na $10 \times 0,05$ j a rovnice funkce intenzity vychází $\lambda(x, y) = 64 \cdot x + 32$.



Obr. 3.1: Cílový průběh funkce délkové intenzity

Definujeme-li délkovou intenzitu v okně pozorování nastavením minimální a maximální intenzity λ_{min} a λ_{max} (čili neznáme střední počet vláken ve čtverci pozorování), pak teoretickou funkci intenzity vyjádříme takto:

$$\lambda(x, y) = (\lambda_{max} - \lambda_{min}) \cdot x + \lambda_{min}. \quad (3.8)$$

Střední počet vláken n_f v okně pozorování lze získat dosazením do vzorce:

$$n_f = \frac{\lambda_{min} + \lambda_{max}}{2 \cdot n_s \cdot l_s}, \quad (3.9)$$

kde n_s je počet segmentů každého vlákna a l_s pevná délka segmentů. U náhodné délky úseček bychom místo deterministického l_s dosadili jejich střední délku.

3.4 Přesnost odhadu funkce intenzity

Přesnost odhadu funkce délkové intenzity reprezentovanou rovinou minimálního kontrastu $\lambda_{mc}(x, y)$ určíme u konkrétního vzorku konfigurace vláken výpočtem odchylek (chyb) od průběhu teoretické funkce intenzity $\lambda(x, y)$. Tímto způsobem je možné vzájemně porovnávat vhodnost příslušných sad nastavení vstupních parametrů.

Absolutní integrovanou chybu $MCIE_a$ (*minimal contrast integrated error*) vypočítáme dvojnou integrací absolutní hodnoty rozdílu roviny minimálního kontrastu $\lambda_{mc}(x, y)$ od teoretické roviny intenzity $\lambda(x, y)$ přes okno pozorování:

$$MCIE_a = \iint_W |\lambda(x, y) - \lambda_{mc}(x, y)| dx dy. \quad (3.10)$$

Protože se však chystáme porovnávat i přesnost odhadů s nestejnou střední hodnotou intenzity v okně, budeme potřebovat též relativní integrovanou odchylku $MCIE_r$ předepsanou zlomkem:

$$MCIE_r = \frac{MCIE_a}{\iint_W \lambda(x, y) dx dy}, \quad (3.11)$$

ve kterém jmenovatel $\iint_W \lambda(x, y) dx dy$ značí teoretickou míru intenzity (integrace cílové funkce intenzity). Násobíme-li stem, obdržíme výsledek v procentech.

Pro navolenou konfiguraci vstupních parametrů vždy realizujeme sto dílčích simulacních experimentů. Z partikulárních chyb zpracovávané dávky určíme aritmetický průměr a výběrovou směrodatnou odchylku. Výsledné hodnoty μ a σ zaznamenáváme do souhrnných tabulek a vykresluje vhodným typem grafu. Nedílnou součástí analýzy jsou i ukázky spektrálních map vybraných vzorků.

4 Simulační experimenty

Pakliže máme definovaný matematický model a naprogramované simulační prostředí, můžeme směle přejít k vlastní realizaci počítačových simulací. Celkem bylo uskutečněno a statisticky zpracováno zhruba 8 000 experimentů, což je pro metodu Monte Carlo plně dostačující. V jejich průběhu byly postupně měněny jednotlivé vstupní parametry a po aplikaci jádrového odhadu vyhodnocována absolutní/relativní odchylka (integrovaná chyba) roviny s minimálním kontrastem od teoretické roviny délkové intenzity.

Nastavované parametry se týkaly jak vlastního procesu generování vláken, tak i stran charakteristik jádrového odhadu a metody minimalizace kontrastu. Konkrétně byl prověřován faktor středního počtu vláken, maximální výchylka, délka a rozlišení segmentu, granularita jádrového odhadu, druh jádrové funkce, šířka pásma, vliv exponentu minimalizace a na samotný závěr aktivace (*post-processing*) eliminace okrajových efektů. Nedílnou součástí analýzy partikulárních faktorů jsou též ukázky spektrálních map pro snadné vizuální posouzení mocnosti vlivu na jádrový odhad.

4.1 Střední počet vláken

Zabývejme se přesností odhadu funkce intenzity (metodou minimálního kontrastu) v závislosti na středním počtu mateřských bodů (prakticky vláken) objevivších se v okně pozorování. Se zvyšujícím se počtem generovaných vláken přirozeně roste i délková intenzita zkoumaného vzorku.

Neměnné složky nastavení simulačního prostředí jsou následující:

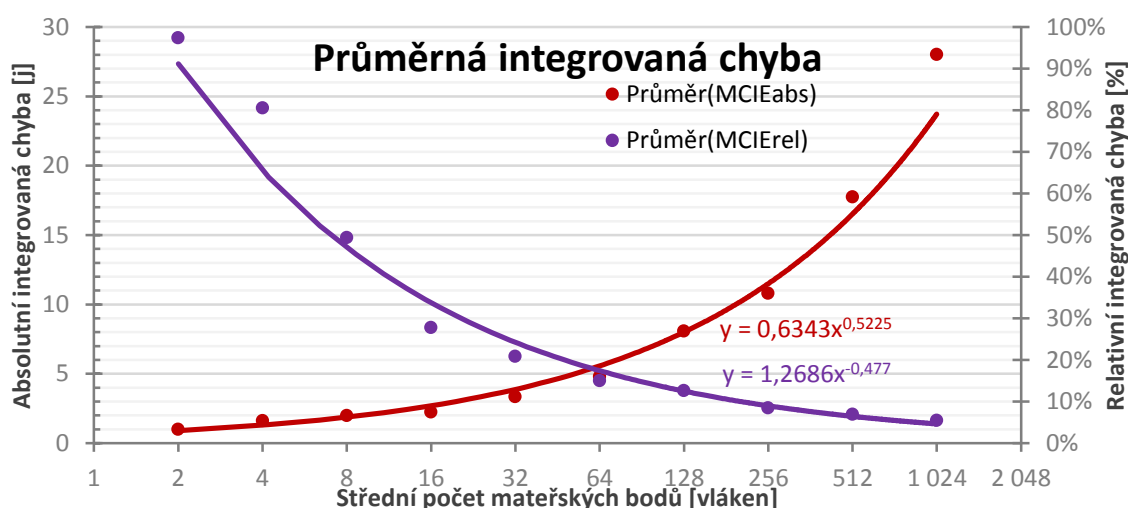
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Pro každý navolený počet vláken (záměrně mocniny čísla 2 – logaritmická osa x) bylo provedeno celkem 100 dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových

směrodatných odchylek, uvádí tabulka 4.1 a názorně vykresluje graf 4.1. Spektrální mapy pro vybrané vzorky s počtem 4, 8, 16, 32, 128 a 512 vláken lze porovnat na obrázku 4.1.

Tab. 4.1: Faktor středního počtu vláken

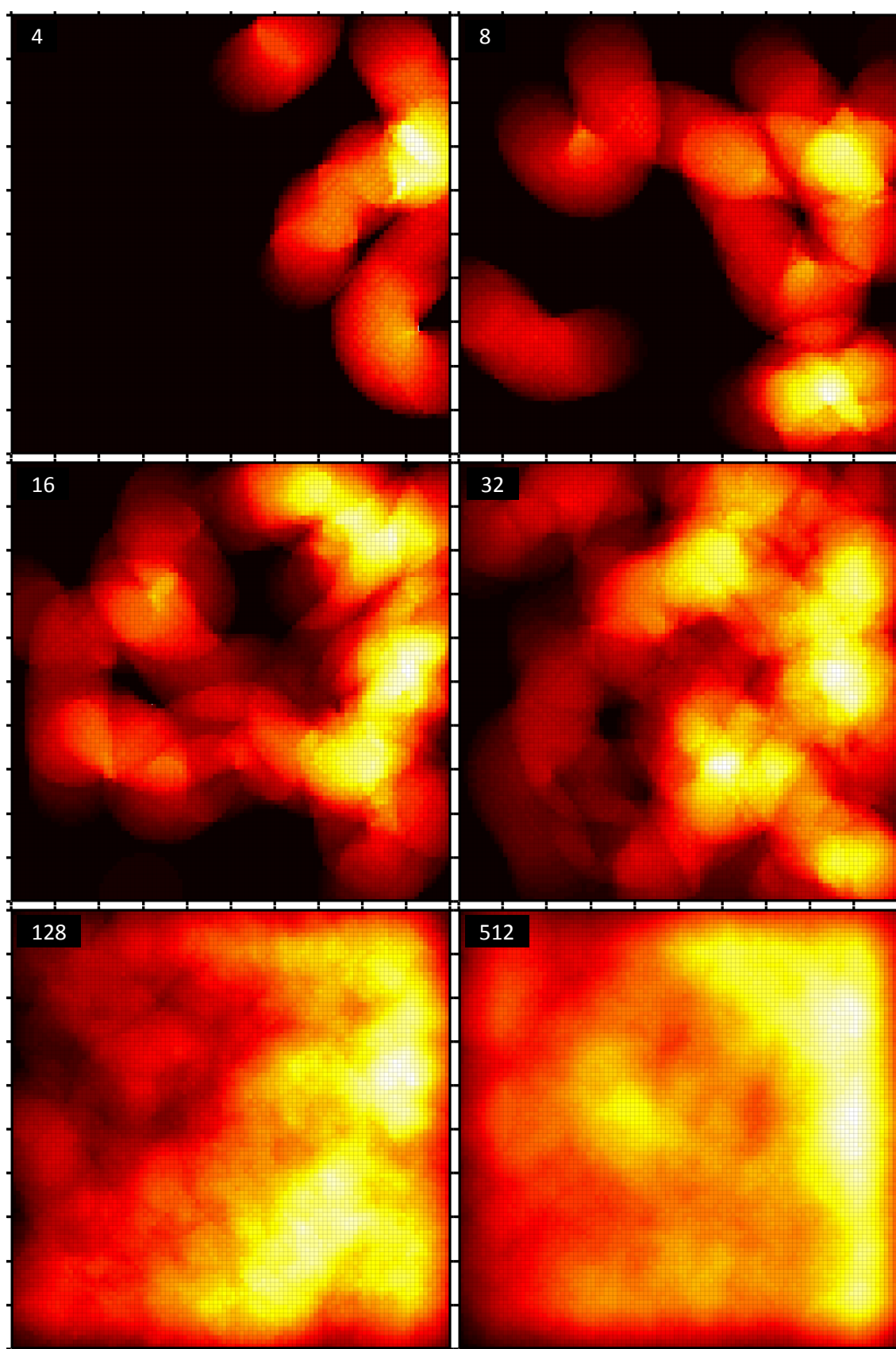
Střední počet vláken	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
2	0,9737	0,1231	97,37%	12,31%
4	1,6091	0,4690	80,45%	23,45%
8	1,9727	0,8084	49,32%	20,21%
16	2,2211	0,7606	27,76%	9,51%
32	3,3278	1,2487	20,80%	7,80%
64	4,7821	2,0543	14,94%	6,42%
128	8,0610	3,4133	12,60%	5,33%
256	10,7962	4,1052	8,44%	3,21%
512	17,7408	7,7341	6,93%	3,02%
1 024	28,0039	11,2854	5,47%	2,20%



Graf 4.1: Faktor středního počtu vláken

Růst absolutní integrované chyby (červená sada hodnot) je zde jevem přirozeným, očekávaným. Relevantní relativní integrovaná chyba (fialová barva) s rostoucím středním počtem vláknem rapidně klesá – aproximací je část křivky mocninné funkce. Jednoznačně tedy můžeme říci, že čím vyšší počet vláken se nachází v okně pozorování (tedy čím vyšší je míra délkové intenzity zkoumaného vzorku), tím přesnější získáme odhad funkce intenzity.

Chceme-li, aby průměrná relativní chyba roviny minimálního kontrastu nepřesáhla 25 %, je nutné pracovat alespoň s 30 vlákny v okně pozorování. Při pouhých deseti instancích již atakujeme hranici 40% relativní odchylky.



Obr. 4.1: Faktor středního počtu vláken

4.2 Maximální výchylka segmentu

Zaměřme pozornost na přesnost odhadu funkce intenzity (metodou minimálního kontrastu) v závislosti na maximální výchylce jednotlivých segmentů v rámci téhož vlákna. Úhel je generován náhodně – v rámci algoritmu realizace náhodné procházky, a to s rovnoměrným rozdělením. Prakticky tedy omezuje hodnotu amplitudy úhlu „na obě strany“ vzhledem předcházejícímu segmentu.

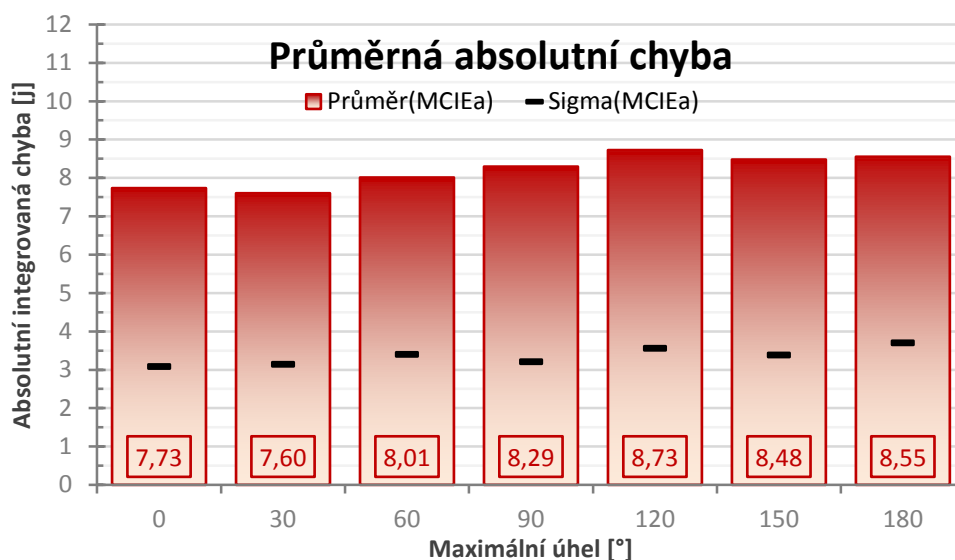
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Pro každou navolenou maximální výchylku bylo provedeno celkem 100 dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.2. Do grafu 4.2 postačí vykreslit odchylky absolutní; střední počet vláken a jejich konfigurace (počet segmentů a jejich délka) zůstávají pro všechny výchylky konstantní. Spektrální mapy pro vybrané vzorky s maximálními výchylkami 30° , 90° a 150° lze porovnat na obrázku 4.2.

Tab. 4.2: Faktor maximální výchylky segmentu

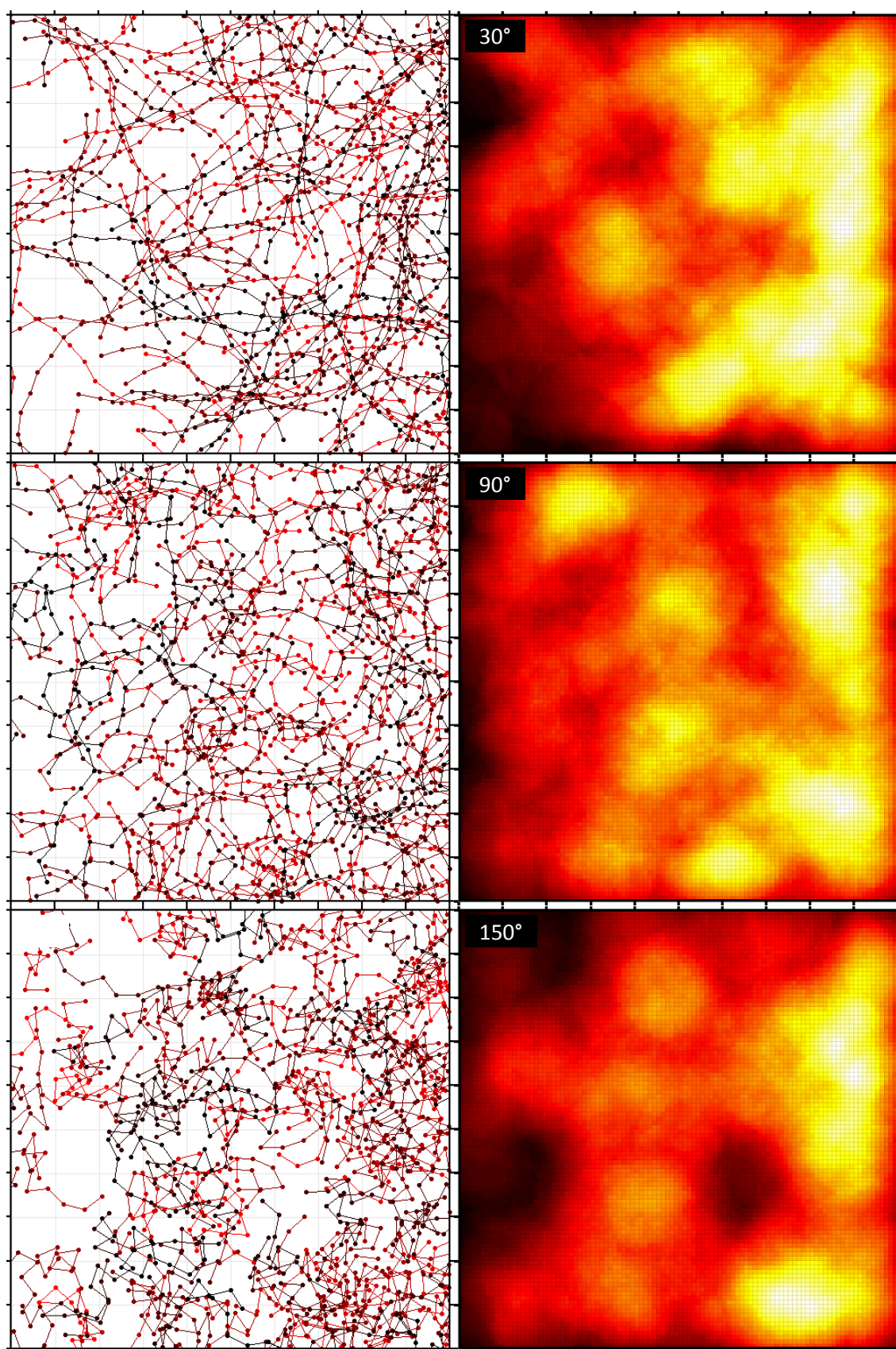
Maximální úhel [°]	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
0	7,7334	3,0769	12,08%	4,81%
30	7,6008	3,1356	11,88%	4,90%
60	8,0103	3,3954	12,52%	5,31%
90	8,2945	3,2004	12,96%	5,00%
120	8,7252	3,5535	13,63%	5,55%
150	8,4785	3,3786	13,25%	5,28%
180	8,5518	3,6935	13,36%	5,77%



Graf 4.2: Faktor maximální výchylky segmentu

Průměrná absolutní chyba se s rostoucí amplitudou úhlu segmentu zvyšuje jen velmi mírně (zároveň pozvolně roste i směrodatná odchylka). Vlákná mají s většími mezními úhly tendenci se „zacuchávat“ – vytváří artefakty (zámotky) a takovéto struktury izolované povahy ovlivňují výsledný výpočet minimalizace kontrastu.

Maximální úhlovou výchylku segmentu si pro následující simulační experimenty ponecháme nastavenou na 60°. Vyhneme se tím smotávání vláken do sebe sama, což by mohlo negativně ovlivnit přesnost metody minimalizace kontrastu; navíc se tím více přiblížíme aplikaci jádrového odhadu nad reálnými daty (například analýza netkaných textilií, nanovláken).



Obr. 4.2: Faktor maximální výchylky segmentu

4.3 Délka segmentu

Dalším zkoumaným faktorem z kategorie způsobu generování vláken je vliv délky jednotlivých segmentů (úseček), ze kterých se vlastní vlákna skládají. Segmenty mohou mít vesměs pevně nastavenou délku, případně je velikost všech segmentů v okně pozorování generována náhodně, a to s rovnoměrným rozdělením $U(0, MAX)$, kde maximální možnou délku segmentu volíme podle momentální potřeby.

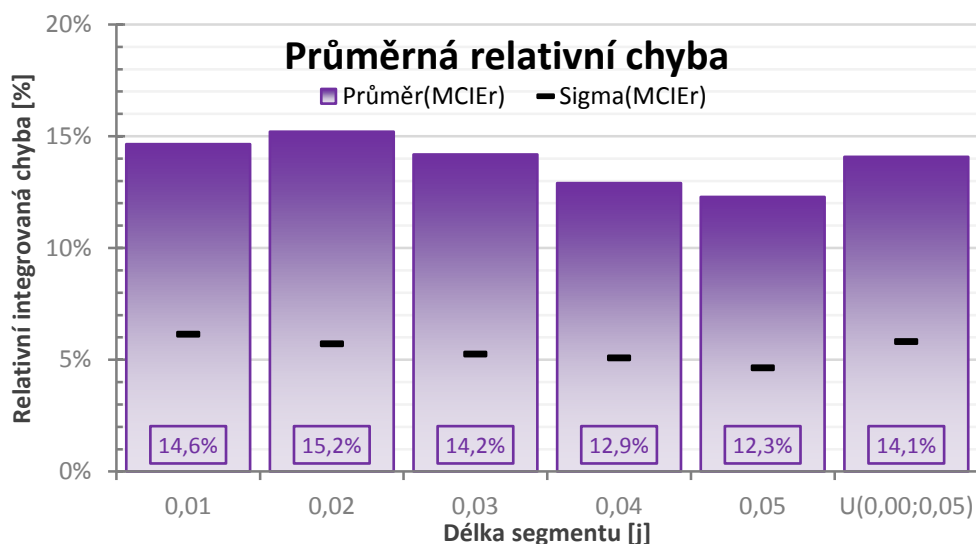
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Postupně byly nastaveny délky 0,01, 0,02, 0,03, 0,04, 0,05 a pro zajímavost i náhodně generovaná velikost segmentu s rozdělením $U(0,00; 0,05)$. Každá konfigurace podstoupila celkem 100 dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.3. Graf 4.3 vykresluje pouze průměrné relativní odchylky; absolutní odchylka přirozeně roste v přímé úměrnosti s rostoucí délkovou intenzitou v okně pozorování. Spektrální mapy pro vybrané vzorky s délkami segmentů 0,02, 0,04 a $U(0,00; 0,05)$ lze vizuálně srovnat na obrázku 4.3.

Tab. 4.3: Faktor délky segmentu

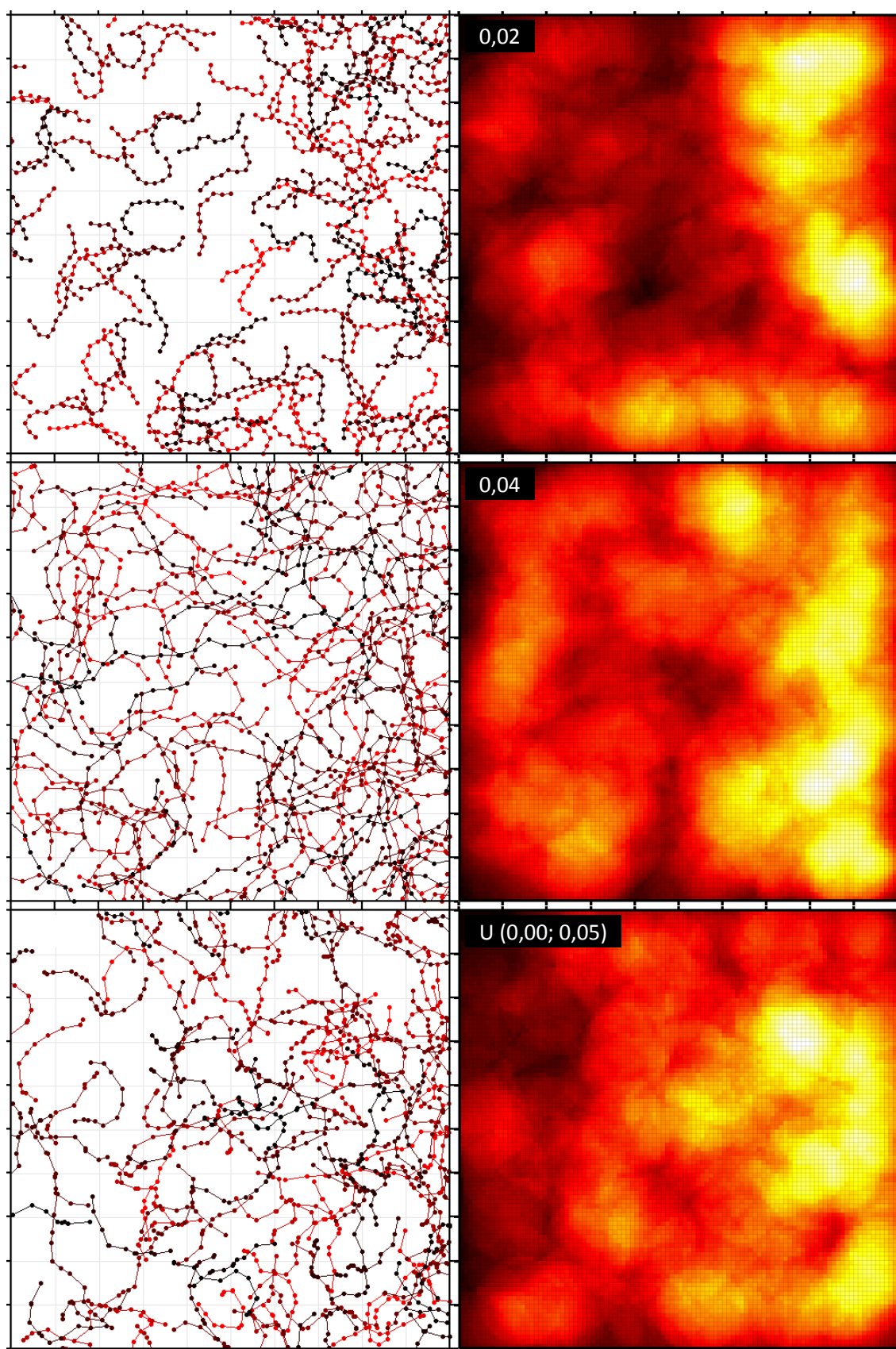
Délka segmentu	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
0,01	1,8743	0,7843	14,64%	6,13%
0,02	3,8912	1,4596	15,20%	5,70%
0,03	5,4457	2,0112	14,18%	5,24%
0,04	6,6046	2,5982	12,90%	5,07%
0,05	7,8595	2,9627	12,28%	4,63%
U (0,00; 0,05)	4,5041	1,8542	14,08%	5,79%



Graf 4.3: Faktor délky segmentu

Průměrná relativní chyba se s rostoucí délkou segmentu snižuje, avšak řádově pouze o jednotky procent (zároveň pozvolně klesá i směrodatná odchylka). Tento jev je v dokonalém souladu s již prověřenou skutečností, kdy s rostoucí délkovou intenzitou v okně pozorování roste i přesnost samotného odhadu. Chyba náhodně generovaných délek úseček přibližně odpovídá průměrné odchylce střední hodnoty příslušného rovnoměrného rozdělení (zde konkrétně 0,025).

Můžeme tedy prohlásit, že samotná délka segmentu nemá zásadní vliv na přesnost odhadu funkce intenzity metodou minimálního kontrastu. V drtivé většině případů si velikost segmentu pro následující simulační experimenty ponecháme pevně nastavenou na hodnotě 0,05.



Obr. 4.3: Faktor délky segmentu

4.4 Rozlišení segmentu

Plynule přecházíme do druhé, patrně i zajímavější oblasti, která se týká konfigurace vlastního jádrového odhadu. Rozlišením segmentu rozumíme skutečnost, s jakou jemností rozmístíme na jednotlivé úsečky vláken rovnoměrně body, do kterých pak umístíme středy jader (respektive jádrových funkcí). Kupříkladu rozlišení „1“ tedy znamená, že každý segment reprezentují dvě jádra – jedno se středem v počátečním a druhé v koncovém bodě úsečky; granularita (jemnost) má hodnotu délky příslušného segmentu. Rozlišení „2“ dělí segment na dva sektory a jádra tedy aplikujeme celkem tři; granularita je rovna polovině velikosti segmentu. A takto bychom mohli pokračovat dále.

Dělením segmentů na drobnější úseky vzniká nutnost převážít jádrovou funkci průměrnou velikostí úsečky, kterou příslušené jádro zastupuje (jejíž „hmotu“ reprezentuje, rozměňuje). To se vztahuje i na počáteční a koncové body celých vláken (za předpokladu tolerance drobné nepřesnosti). Délku celého vlákna tedy vydělíme počtem jader, která jej aproximují, a výslednou hodnotou korigujeme všechny jádrové funkce obsažené v daném vlákně.

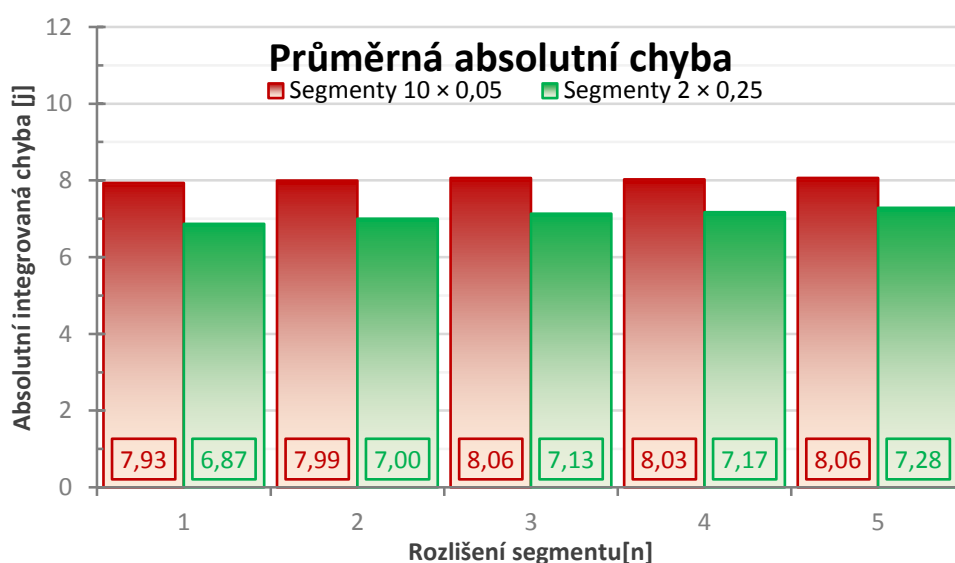
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $2 \times 0,25$ j a $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Prakticky bylo vybráno rozlišení segmentů 1, 2, 3, 4 a 5; jemněji již nemělo smysl pokračovat. Pro každou takto navolenou granularitu bylo provedeno celkem 100 dílčích experimentů. Sada simulačních pokusů navíc proběhla zvlášť s dvěma různými konfiguracemi vláken shodné délky – konkrétně $2 \times 0,25$ j a $10 \times 0,05$ j. Výčet průměrných hodnot absolutních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.4. Graf 4.4 vykresluje průměrné absolutní odchylky, a to pro obě konfigurace vláken. Spektrální mapy vybraných vzorků vláken typu $10 \times 0,05$ j s rozlišením 1 až 5 lze vizuálně porovnat na obrázku 4.4.

Tab. 4.4: Faktor rozlišení segmentu

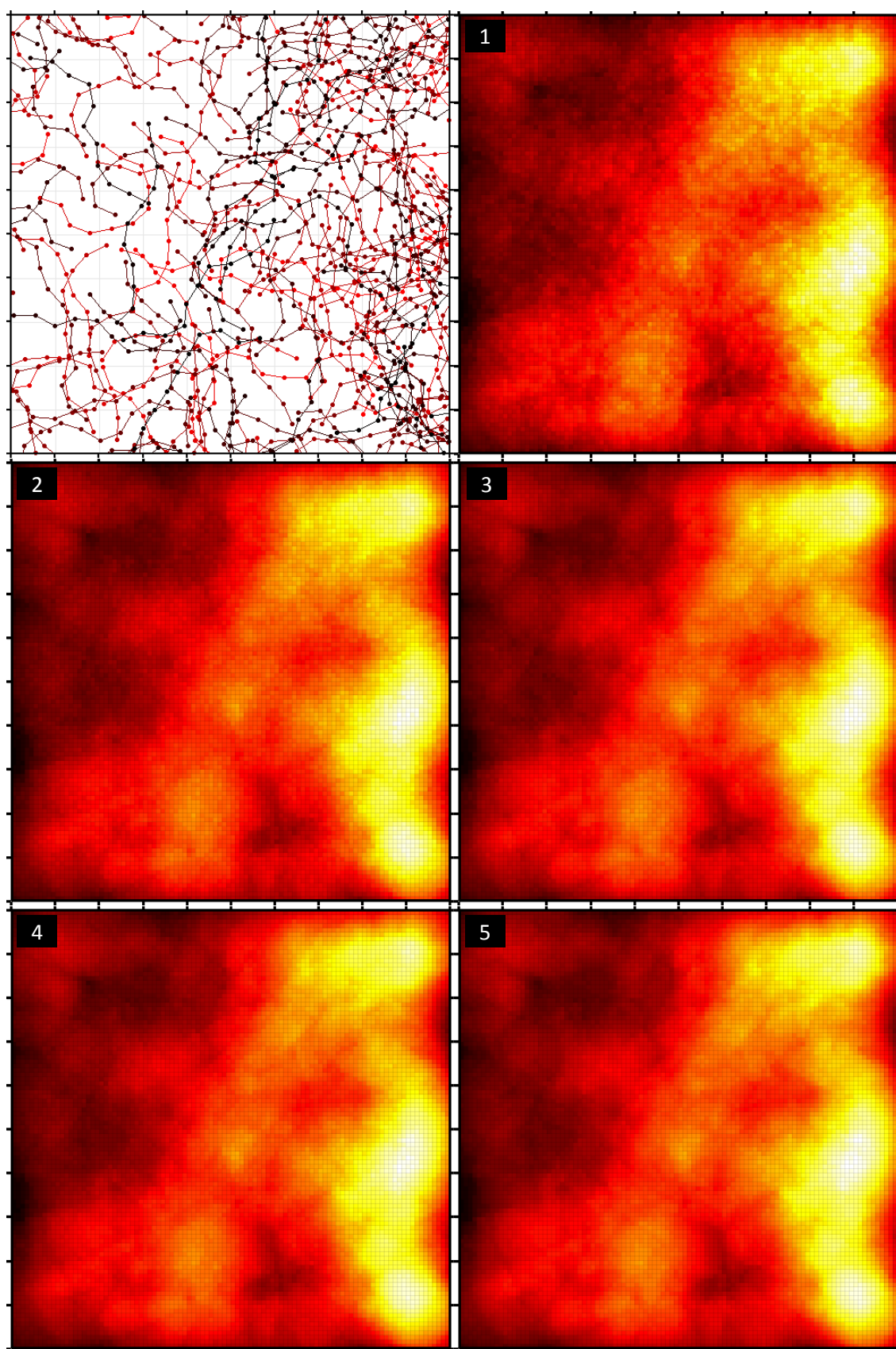
Rozlišení segmentu	Segmenty $2 \times 0,25$		Segmenty $10 \times 0,05$	
	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
1	6,8665	3,2703	7,9326	2,9475
2	7,0008	3,3362	7,9949	2,9931
3	7,1311	3,3673	8,0619	3,0031
4	7,1701	3,3697	8,0259	3,0669
5	7,2844	3,3504	8,0633	3,0118



Graf 4.4: Faktor rozlišení segmentu

Absolutní chyba (i její výběrová směrodatná odchylka) je pro obě konfigurace vláken víceméně konstantní. Zjemňováním granularity rozlišení segmentu odhad intenzity metodou minimálního kontrastu prakticky nezpřesníme. Nadále si ponecháme pevně nastavenou (značně konzervativní) hodnotu rozlišení úseček na čísle 4 – pro segmenty délky 0,05 j tedy vychází granularita 0,0125 j.

Za zmínku stojí zřejmě nižší průměrná odchylka v případě vlákenné konfigurace $2 \times 0,25$ j vůči rozmístění $10 \times 0,05$ j. Zdánlivě překvapivý jev lze vysvětlit tím, že v případě rozdělení vlákna na pouhé dva dlouhé segmenty dochází k mnohem rovnoměrnějšímu rozložení jader (podél dlouhých úseček), a tedy přesnějšímu jádrovému odhadu. Praktická aplikace metody při analýze netkaných textilií nás každopádně vede spíše k opačnému rozložení (jemnějšímu vzorkování/bodování jednotlivých vláken).



Obr. 4.4: Faktor rozlišení segmentu

4.5 Rozlišení jádrového odhadu

V tomto pododdílu zjišťujeme mocnost vlivu výstupní granularity, zrnitosti. Rozlišení jádrového odhadu si můžeme představit jako velikost výstupní matice uchovávající hodnoty jádrového odhadu, jejíž vlastnosti dědí i obrazové rozlišení spektrálních map v „pixelech“, obrazových bodech. Přesněji řečeno jde o „jemnost“ kroku numerické integrace jádrových funkcí.

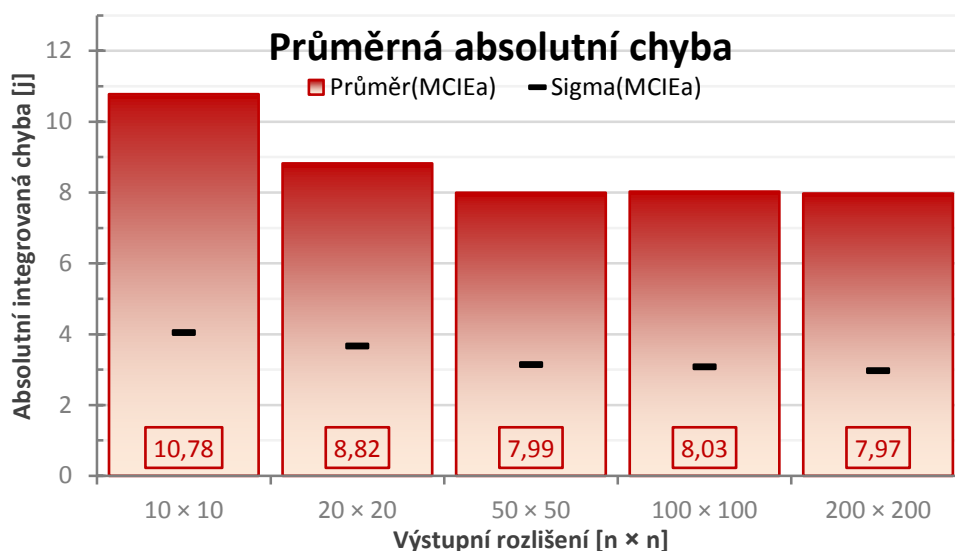
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Konkrétně byla vybrána výstupní rozlišení 10×10 , 20×20 , 50×50 , 100×100 a 200×200 . Pro každou z těchto zrnitostí bylo provedeno celkem 100 dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.5. Průměrné absolutní chyby zobrazuje graf 4.5. Spektrální mapy vybraných vzorků s rozlišením 10×10 až 200×200 lze vizuálně komparovat na obrázku 4.5.

Tab. 4.5: Faktor rozlišení jádrového odhadu

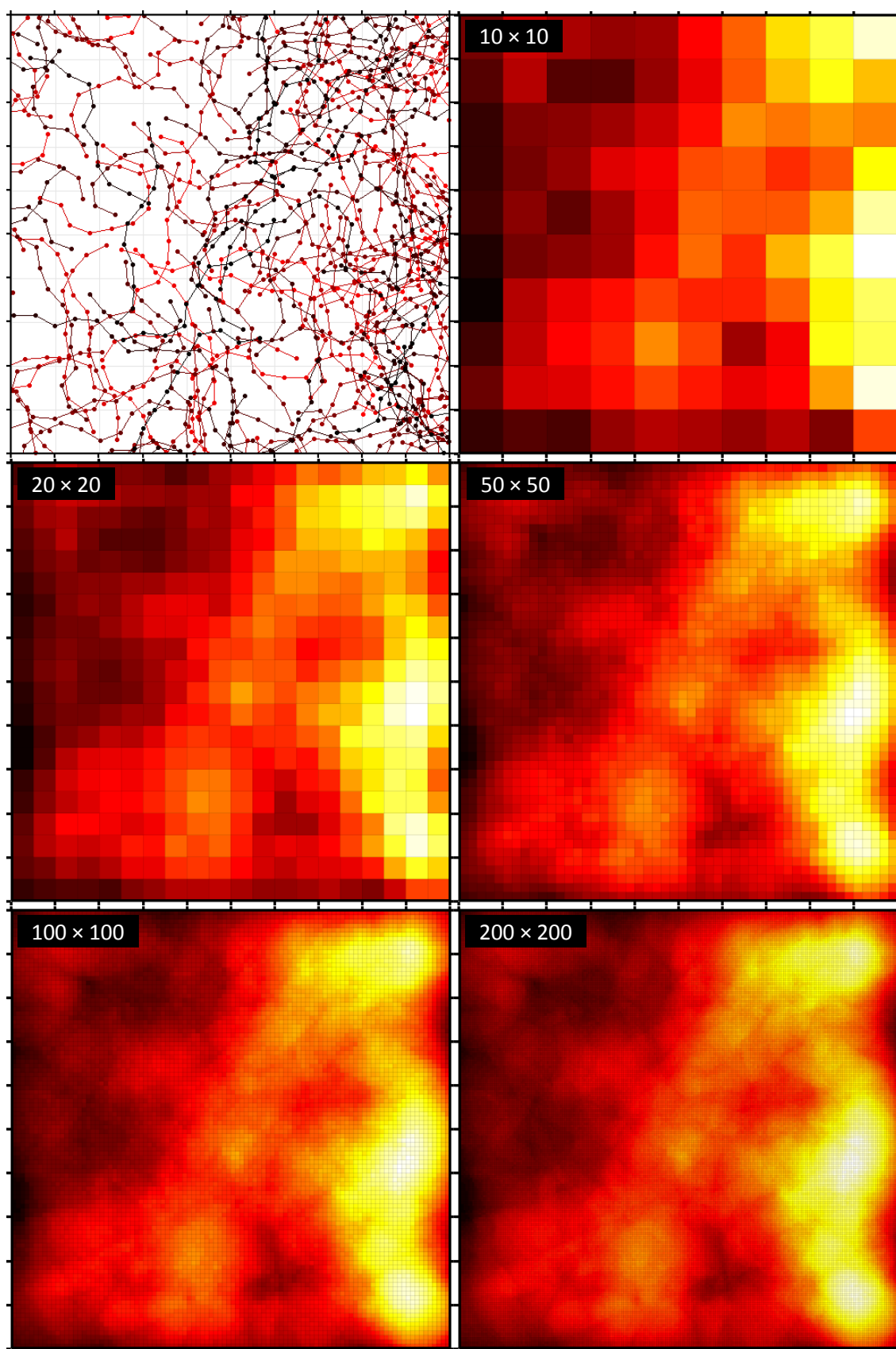
Výstupní rozlišení	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
10 × 10	10,7761	4,0362	16,84%	6,31%
20 × 20	8,8217	3,6608	13,78%	5,72%
50 × 50	7,9941	3,1345	12,49%	4,90%
100 × 100	8,0259	3,0669	12,54%	4,79%
200 × 200	7,9716	2,9625	12,46%	4,63%



Graf 4.5: Faktor rozlišení jádrového odhadu

Přesně podle očekávání můžeme konstatovat, že čím „jemněji“ integrujeme jádrový odhad, tím je tento přesnější, což má pochopitelně pozitivní účinek na minimalizaci kontrastu. Vyšší stupně rozlišení ovšem vyžadují (násobně) větší rozměr příslušné matice a zpracování bývá výpočetně mnohem náročnější – jak na operační paměť, tak i přirozeně na dobu výpočtu, která roste exponenciálně.

Obecně lze doporučit minimální hodnotu rozlišení 50×50 (odpovídá granularitě $0,02 \text{ j}^2$). Jemnější zrnitostí rozhodně nic „nezkazíme“, ale absolutní integrovaná chyba již dále nijak významně neklesá; stejně tak ani výběrová směrodatná odchylka. Na spektrálních mapách s výstupním rozlišením 100×100 (granularita $0,01 \text{ j}^2$) a 200×200 (zrnitost $0,005 \text{ j}^2$) již nejsou patrné žádné větší rozdíly; k obrazovým artefaktům nedochází.



Obr. 4.5: Faktor rozlišení jádrového odhadu

4.6 Jádrová funkce

Dalším zkoumaným faktorem stran konfigurace jádrového odhadu je mocnost vlivu druhu použitého jádra. Přesvědčíme se, že volba jádrové funkce nemá kruciální dopad na odhad funkce intenzity metodou minimálního kontrastu.

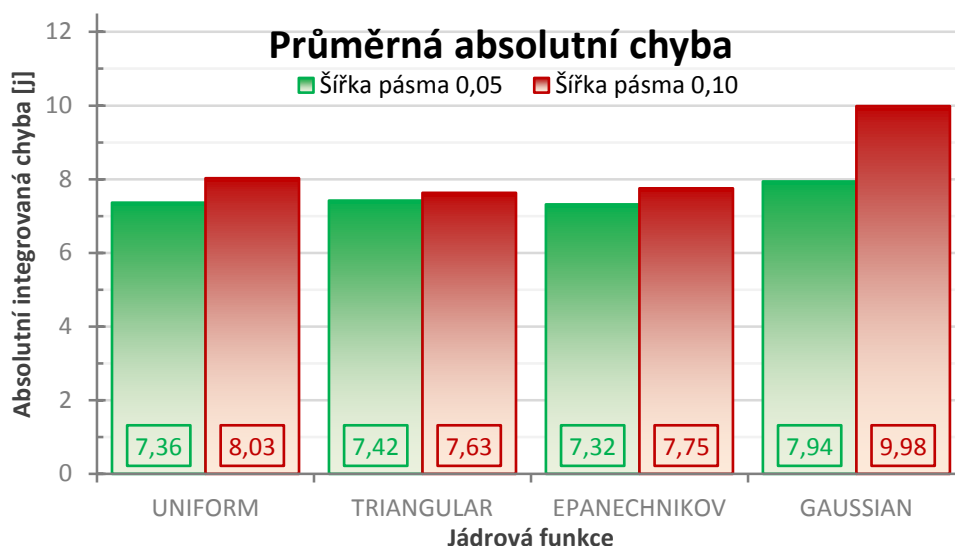
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

V naprogramovaném simulačním prostředí máme k dispozici všechna čtyři již dříve představená a podrobně charakterizovaná jádra: indikátor okolí bodu (*UNIFORM*), lineární převážení (*TRIANGULAR*), kvadratické převážení (*EPANECHNIKOV*) a normální převážení (*GAUSSIAN*). Každý z těchto typů jádrových funkcí podstoupil celkem 100 dílčích experimentů. Stejná sada simulačních vzorků navíc posloužila zvláště pro dvě různé šířky pásma – konkrétně 0,05 j a 0,10 j. Výčet průměrných hodnot absolutních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.6. Graf 4.6 vykresluje průměrné absolutní odchylky, a to pro obě konfigurace šířky pásma. Spektrální mapy vybraných vzorků jsou k nahlédnutí na obrázku 4.6; pásmo zde má šířku 0,10 j.

Tab. 4.6: Faktor druhu jádrové funkce

Jádrová funkce	Šířka pásma 0,05		Šířka pásma 0,10	
	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
UNIFORM	7,3627	2,8011	8,0259	3,0669
TRIANGULAR	7,4188	2,9247	7,6294	2,8480
EPANECHNIKOV	7,3158	2,9520	7,7516	2,8403
GAUSSIAN	7,9384	2,9511	9,9839	3,5397

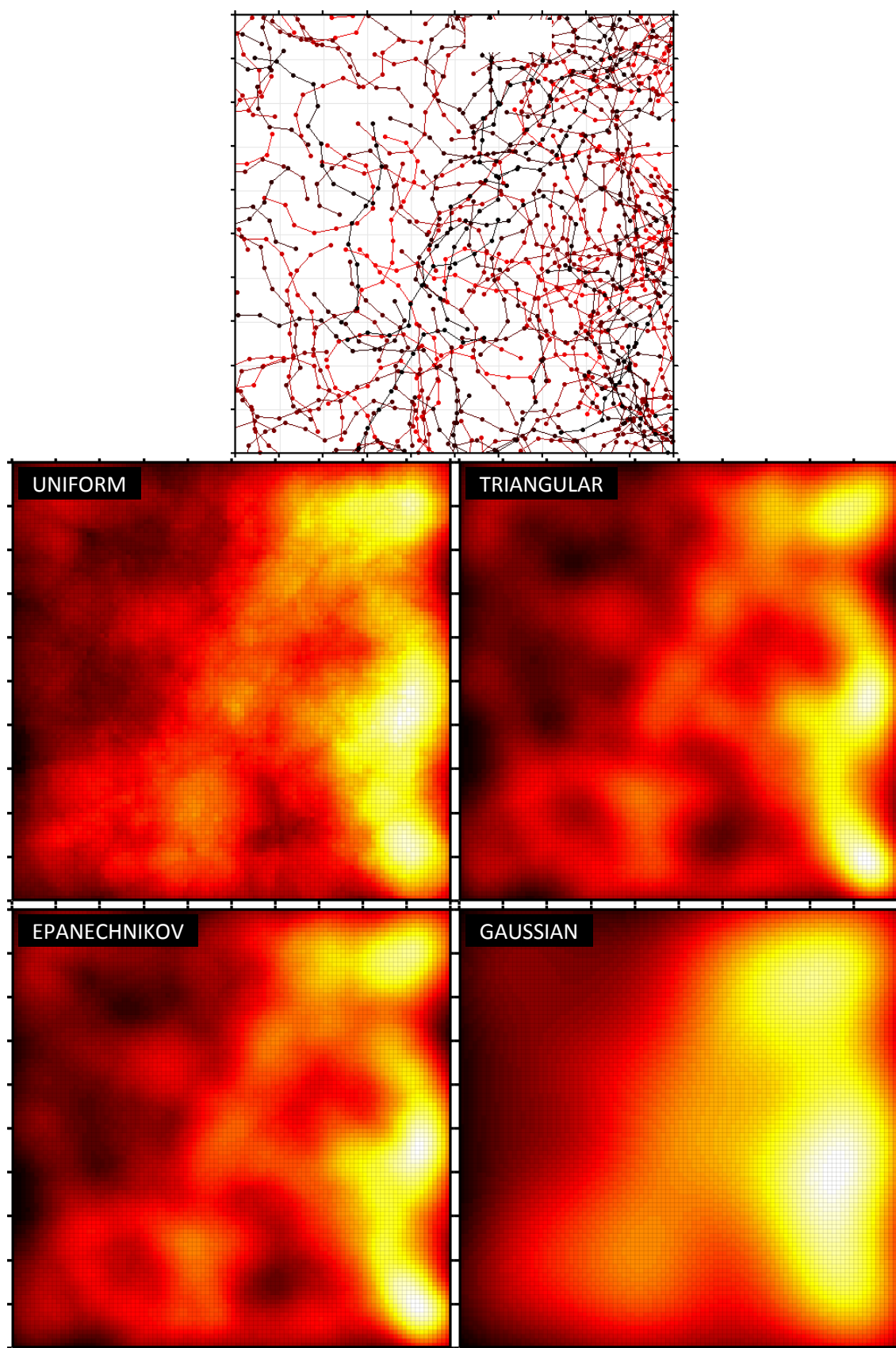


Graf 4.6: Faktor druhu jádrové funkce

Indikátor okolí bodu, lineární a kvadratické převážení produkují velmi podobné výsledky minimalizace kontrastu, a to pro obě navolené šířky pásma. Průměrné absolutní chyby nijak výrazně nevybočují, lze doporučit kterékoli z nich. Jádro typu *UNIFORM* vykresluje poněkud ostřejší přechody (vzdáleně připomínající atmosférická oblaka). Oproti tomu spektrální mapy jader *TRIANGULAR* a *EPANECHNIKOV* je od sebe velmi obtížné odlišit – vytváří měkké, jakoby tlumené kontury.

Normální rozdělení vykazuje velkou míru vyhlazení (*smoothing*). Jádro typu *GAUSSIAN* je proto vhodnější používat pouze s užšími šířkami pásma; rozdíl mezi námi testovanými šířkami 0,10 a 0,05 je na první pohled zřejmý. Širší pásmo vede k příliš velkému vyhlazení scény (efekt zvaný *over-smoothing*).

Ve všech čtyřech případech si můžeme povšimnout výrazného vlivu okrajových efektů. Jejich eliminaci zapneme a účinnost vzhledem k minimalizaci kontrastu posoudíme v pododdílu 4.9.



Obr. 4.6: Faktor druhu jádrové funkce

4.7 Šířka pásma

V rámci tohoto pododdílu se soustředíme na správnou volbu šířky pásma (*bandwidth selection*). Jedná o jeden z klíčových parametrů konfigurace jádrového odhadu.

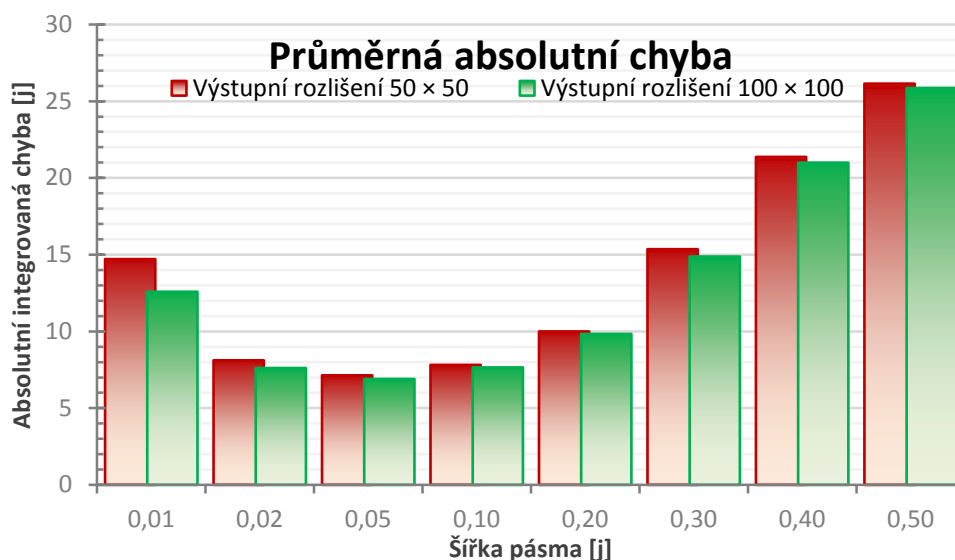
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 50×50 a 100×100
- exponent minimalizace: 1
- vypnutá eliminace okrajových efektů

Postupně otestujeme následující šířky pásma: 0,01, 0,02, 0,05, 0,10, 0,20, 0,30, 0,40 a 0,50. Pro každou z těchto šířek provedeme obligátních 100 dílčích experimentů. Všechny vzorky zpracujeme ve dvou výstupních rozlišeních, 50×50 a 100×100 . Výčet průměrných hodnot absolutních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.7. Graf 4.7 vykresluje průměrné absolutní odchylky, a to pro obě dvě zvolené granularity. Spektrální mapy vybraných vzorků na šířce pásma 0,02, 0,05, 0,10, 0,20 a 0,30 lze vizuálně porovnat na obrázku 4.7; zrnitost je zde $0,01 \text{ j}^2$.

Tab. 4.7: Faktor šířky pásma

Šířka pásma	Výstupní rozlišení 50×50		Výstupní rozlišení 100×100	
	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _a)	Sigma (MCIE _a)
0,01	14,7086	5,9028	12,5846	5,3392
0,02	8,1150	3,3215	7,6124	3,1493
0,05	7,1366	3,1477	6,9015	2,8928
0,10	7,8159	3,2039	7,6523	3,0454
0,20	9,9961	3,6738	9,8326	3,4381
0,30	15,3556	3,1033	14,8781	3,1495
0,40	21,3672	2,8351	20,9942	2,8632
0,50	26,1419	2,6613	25,8576	2,6917

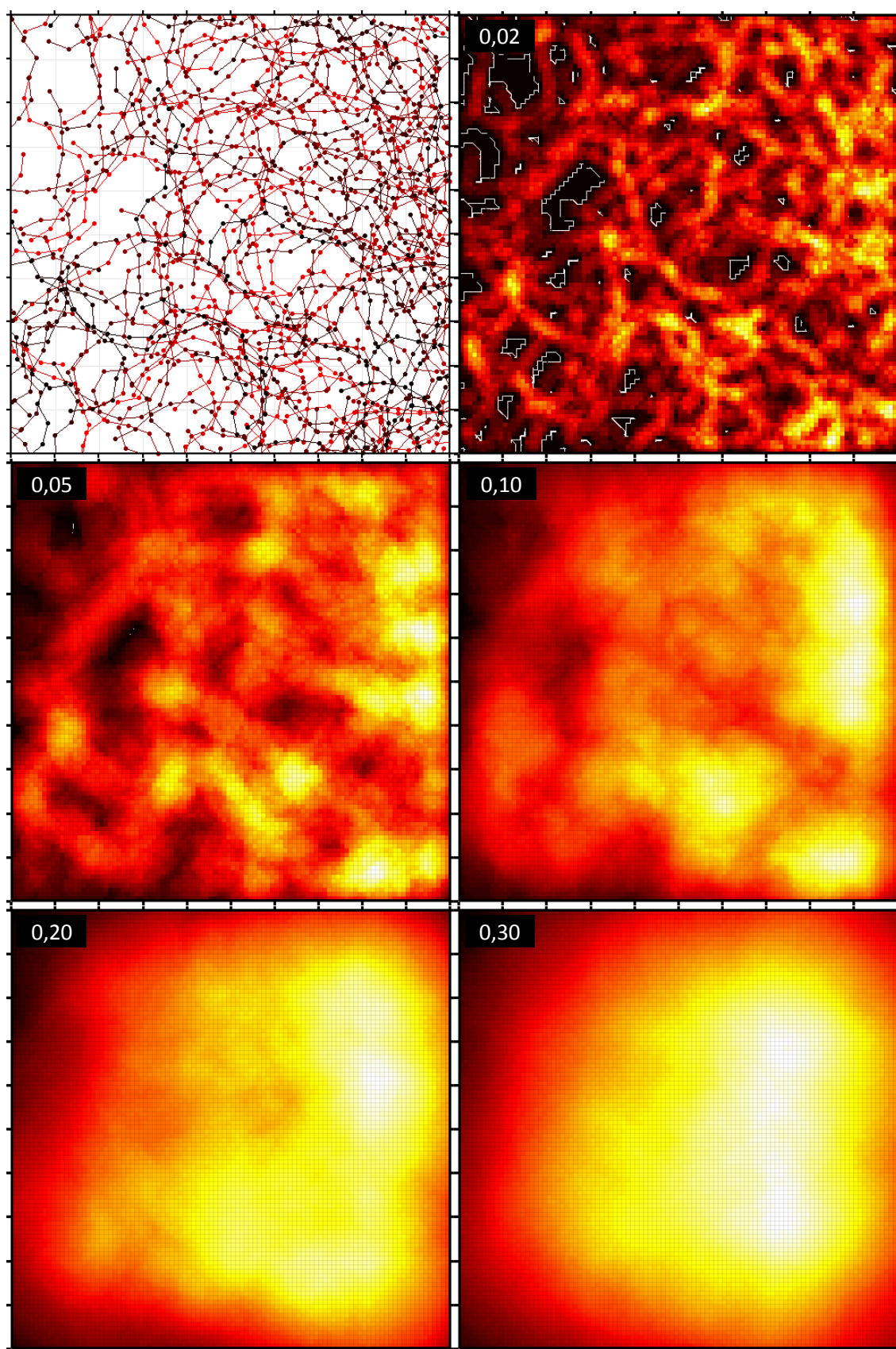


Graf 4.7: Faktor šířky pásma

Úzké pásmo jádrové funkce vede k nízkému stupni vyhlazení (efekt označovaný termínem *under-smoothing*) – spektrální mapa takové konfigurace působí podexponovaně. Naopak širší pásma vytváří vyhlazení příliš velké (jev známý jako *over-smoothing*); odpovídající spektrální mapa je přexponovaná.

Jako ideální se jeví šířky pásma 0,05 a 0,10 – absolutní integrované chyby roviny minimálního kontrastu jsou v těchto konfiguracích nejnižší a totéž platí i pro výběrové směrodatné odchylky. Rozdíl mezi granularitou 0,01 j^2 a 0,02 j^2 je pro takto nastavená pásma nevelký; přirozeně však o něco lépe vyznívá pro rozlišení 100 × 100.

Na první pohled se sice nabízí též šířka pásma o velikosti 0,02, ale po prozkoumání spektrálních map s tímto nastavením zjistíme značně nízkou míru vyhlazení celé scény. Jak již víme z výsledků předchozích simulací, úzká pásma jsou vhodná pouze pro jádrové funkce typu *GAUSSIAN* (normální převážení).



Obr. 4.7: Faktor šířky pásma

4.8 Exponent minimalizace

Při vlastní minimalizaci kontrastu můžeme použít prakticky libovolný (nenulový) přirozený exponent. Zvolíme-li hodnotu 1, minimalizujeme absolutní hodnoty kontrastu *IAE* (*Integrated Absolute Error*), při hodnotě 2 minimalizujeme čtverce kontrastu *ISE* (*Integrated Square Error*) a takto bychom mohli pokračovat do nekonečna. Jak velký vliv má volba exponentu na výslednou rovinu minimálního kontrastu?

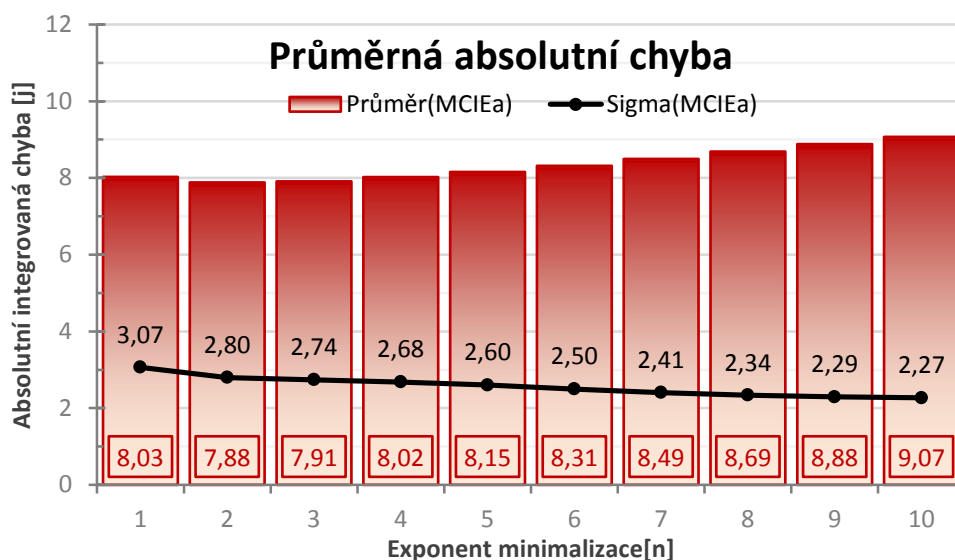
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: indikátor okolí bodu (*UNIFORM*)
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- vypnutá eliminace okrajových efektů

Postupně otestujeme exponenty od 1 do 10. Pro každou z těchto hodnot provedeme obvyklých 100 dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.8. Graf 4.8 vykresluje průměrné absolutní chyby a jejich výběrové směrodatné odchylky. Spektrální mapy si tentokrát neukážeme, protože samotný výpočet kontrastu a jeho minimalizace probíhá až po integraci jádrových funkcí (výstupní matice je proto pro všechny vzorky konstantní).

Tab. 4.8: Faktor exponentu minimalizace

Exponent minimalizace	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _r)	Sigma (MCIE _r)
1	8,0259	3,0669	12,54%	4,79%
2	7,8816	2,7977	12,31%	4,37%
3	7,9093	2,7416	12,36%	4,28%
4	8,0217	2,6818	12,53%	4,19%
5	8,1548	2,6041	12,74%	4,07%
6	8,3147	2,4980	12,99%	3,90%
7	8,4942	2,4059	13,27%	3,76%
8	8,6863	2,3396	13,57%	3,66%
9	8,8805	2,2941	13,88%	3,58%
10	9,0693	2,2671	14,17%	3,54%



Graf 4.8: Faktor exponentu minimalizace

S rostoucím exponentem minimalizace kontrastu pozvolna roste i průměrná absolutní chyba, výběrová směrodatná odchylka nepatrně klesá. V reálných aplikacích si tedy můžeme bez obav vystačit s hodnotou exponentu rovnou jedné (minimalizace absolutní odchylky), případně volíme dvojku (metoda nejmenších čtverců), či trojku. Nastavování vyšších exponentů již smysl prakticky nemá – odchylka roviny minimálního kontrastu narůstá.

4.9 Eliminace okrajových efektů

Problém s okrajovými efekty (termíny *edge effects* a *boundary bias*) řešíme výpočtem individuálních koeficientů (korigujících ztrátu informací nacházejících se mimo okno pozorování) pro každý prvek výstupní matice integrovaných jádrových funkcí. Jedná se tedy o formu *post-processing* efektu, potažmo filtru.

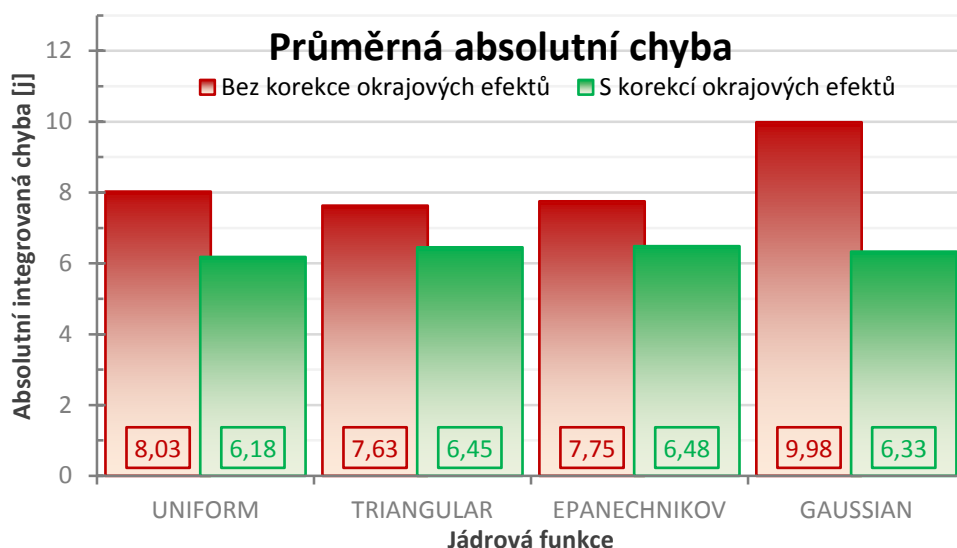
Neměnné složky nastavení simulačního prostředí jsou následující:

- střední počet vláken v $[0,1]^2$: 128
- konfigurace vláken: $10 \times 0,05$ j
- maximální výchylka segmentu: 60°
- jádrová funkce: *UNIFORM*, *TRIANGULAR*, *EPANECHNIKOV* a *GAUSSIAN*
- šířka pásma: 0,1 j
- rozlišení segmentu: 4
- rozlišení jádrového odhadu: 100×100
- exponent minimalizace: 1

Postupně vyhodnotíme vliv okrajových efektů před a po aktivaci jejich eliminace na všechna čtyři připravená jádra. Na každém jádře tedy provedeme $100 + 100$ dílčích experimentů. Výčet průměrných hodnot absolutních a relativních integrovaných chyb roviny minimálního kontrastu, včetně výběrových směrodatných odchylek, uvádí tabulka 4.9. Graf 4.9 vykresluje průměrné absolutní chyby. Spektrální mapy jednotlivých jader před a po aktivaci vidíme na obrázcích 4.9, 4.10, 4.11 a 4.12; konkrétní reprezentativní vzorek vlákené konfigurace zobrazuje obrázek 4.8.

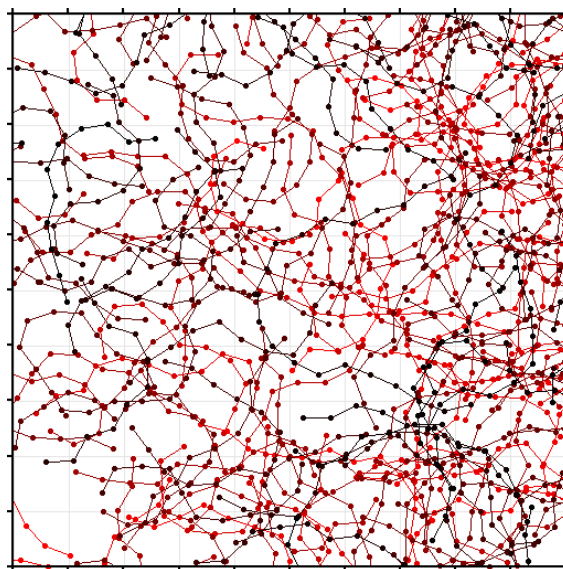
Tab. 4.9: Faktor okrajových efektů

Jádrová funkce	Bez korekce okrajových efektů		S korekcí okrajových efektů	
	Průměr (MCIE _a)	Sigma (MCIE _a)	Průměr (MCIE _a)	Sigma (MCIE _a)
UNIFORM	8,0259	3,0669	6,1803	2,7164
TRIANGULAR	7,6294	2,8480	6,4524	2,8555
EPANECHNIKOV	7,7516	2,8403	6,4848	2,9072
GAUSSIAN	9,9839	3,5397	6,3312	2,5935

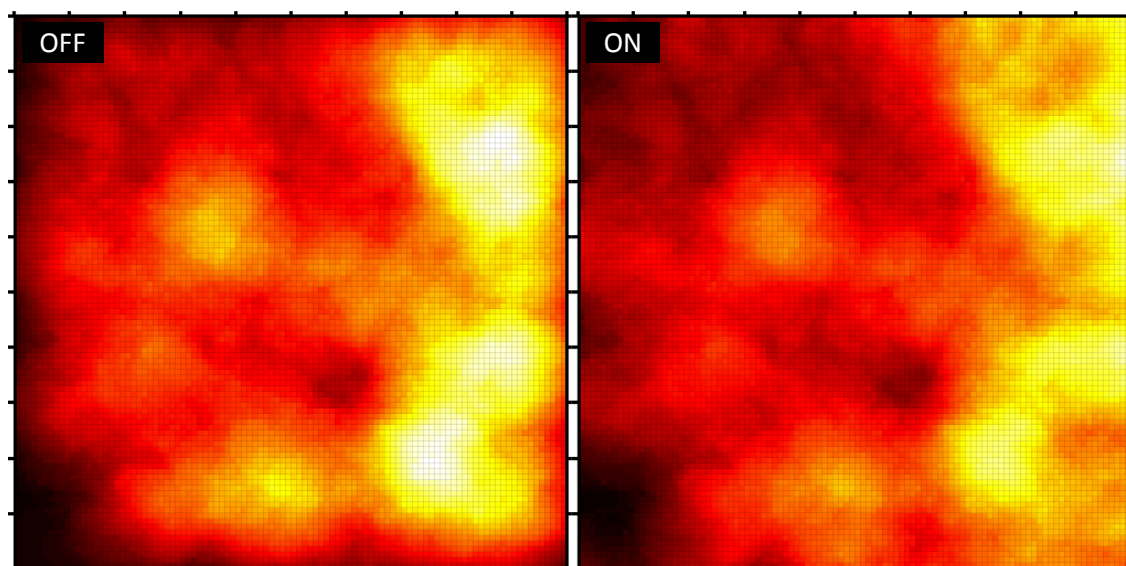


Graf 4.9: Faktor okrajových efektů

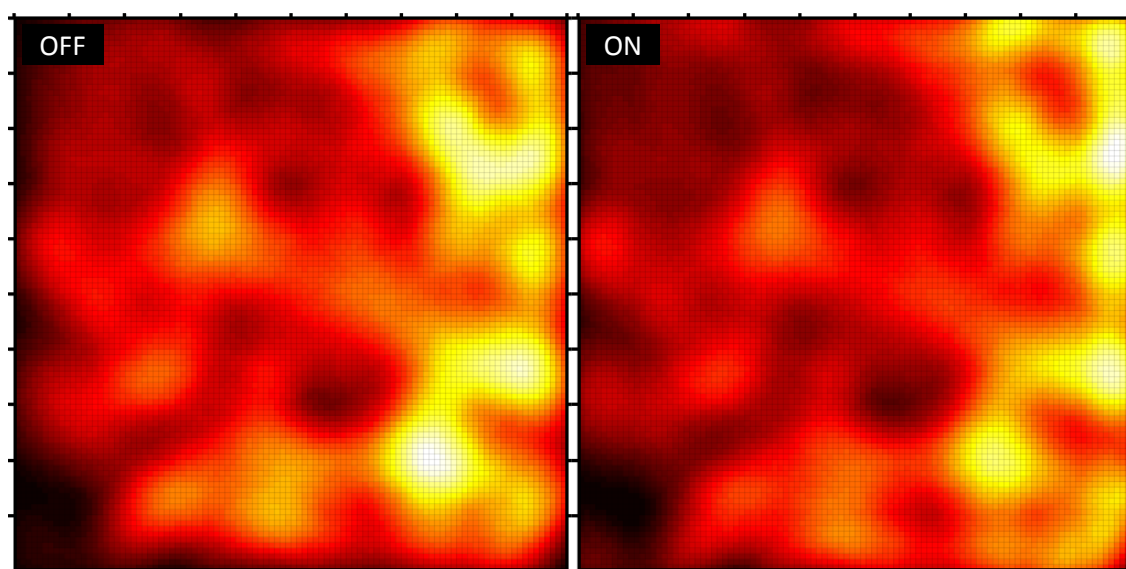
Je naprosto zřejmé, že aktivací eliminace okrajových jevů rapidně klesá průměrná absolutní integrovaná chyba. Zdaleka největšího efektu snížení odchylky roviny minimálního kontrastu se takto dočkáme u normálního převážení, kde se už samotná spektrální mapa s ošetřenými okraji téměř shoduje s rovinou teoretické funkce délkové intenzity.



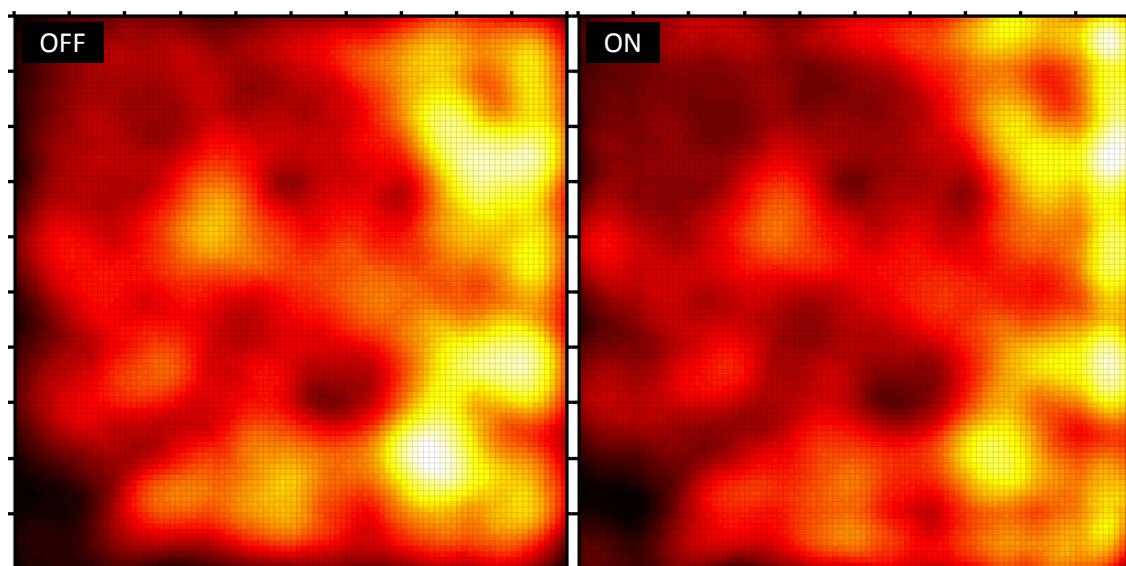
Obr. 4.8: Ukázka analyzované konfigurace



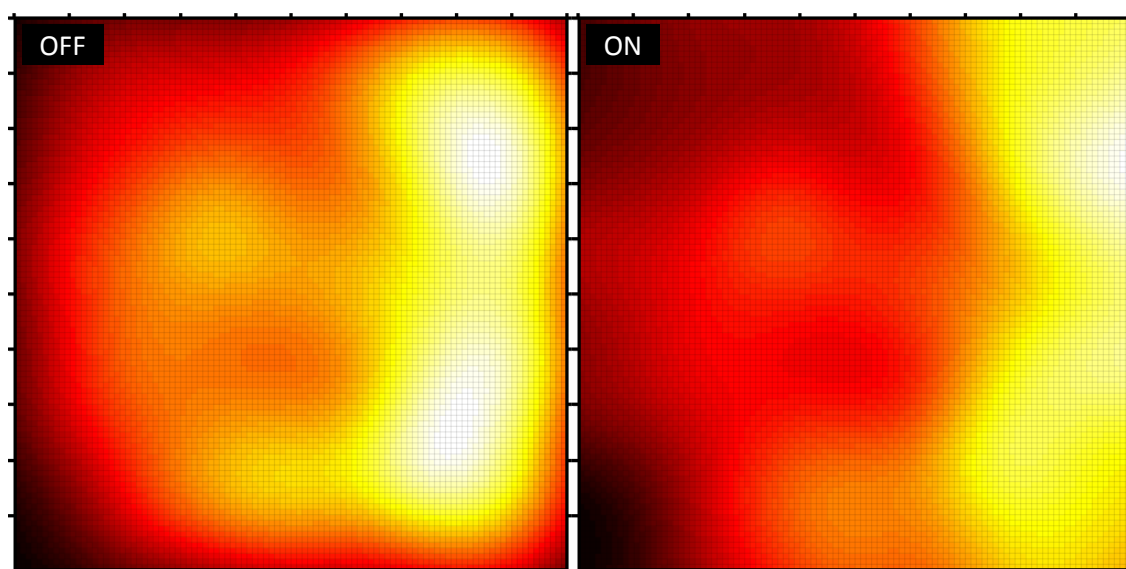
Obr. 4.9: Faktor okrajových efektů (UNIFORM)



Obr. 4.10: Faktor okrajových efektů (TRIANGULAR)



Obr. 4.11: Faktor okrajových efektů (EPANECHNIKOV)



Obr. 4.12: Faktor okrajových efektů (GAUSSIAN)

5 Závěry a doporučení

Bakalářská diplomová práce definovala dostatečně reprezentativní simulační model procesu vláken, u kterého bylo možné přesně určit teoretickou (cílovou) funkci délkové intenzity vláken a stanovit způsoby komparace s rovinou minimálního kontrastu jádrového odhadu. Numerická aproximace zvoleného abstraktního modelu byla naprogramována v softwarovém skriptovacím prostředí MATLAB®. V takto vzniklém virtuálním prostoru proběhlo dostatečné množství simulačních experimentů metody Monte Carlo, a to při postupných změnách jednotlivých vstupních parametrů. Přesnosti odhadu funkce intenzity jednotlivých vzorků, respektive absolutní a relativní chyby rovin minimálního kontrastu integrovaných jádrových funkcí, byly statisticky zpracovány a vzájemně porovnány. Naším cílem bylo nalézt účinné techniky, respektive konkrétní doporučená nastavení, zejména pro jádrový odhad samotný, vedoucí k co možná nejpřesnější aproximaci funkce délkové intenzity.

Zabýváme-li se seskupením a konfigurací vláken v okně pozorování, pak přirozeně potvrzujeme, že s rostoucím středním počtem vláknem odchylka rapidně klesá. Chceme-li, aby průměrná relativní chyba roviny minimálního kontrastu nepřesáhla 25 %, je nutné pracovat alespoň s 30 vlákny. S rostoucí amplitudou úhlu segmentu se chyba zvyšuje jen velmi mírně – vlákna mají s většími mezními úhly tendenci se „zacuchávat“ a takovéto struktury izolované povahy mírně negativně ovlivňují výsledný výpočet minimalizace. Samotná délka segmentu nemá zásadní vliv na přesnost odhadu funkce intenzity metodou minimálního kontrastu.

Ohledně výstupního rozlišení numerické aproximace lze doporučit minimální hodnotu 50×50 (odpovídá granularitě $0,02 \text{ } \mu\text{m}^2$). Jemnější zrnitost rozhodně nic nezkažeme, ovšem integrovaná chyba již dále významně neklesá. Náročnost výpočtu roste geometrickou řadou – záleží tedy i na dostupných výpočetních prostředcích. Jsou-li segmenty jednotlivých vláken „dostatečně krátké“, pak zjemňováním granularity segmentu odhad intenzity metodou minimálního kontrastu prakticky nezpřesníme. Prakticky stačí posoudit partikulární vztah mezi vzdáleností středů jádrových funkcí v rámci příslušné úsečky, šířkou pásma aplikovaného jádra a výstupní granularitou jádrového odhadu.

Simulační prostředí nabízí výběr ze čtyř jádrových funkcí. Indikátor okolí bodu, lineární a kvadratické převážení produkují velmi podobné výsledky; průměrné odchylky nijak výrazně nevybočují a nebojíme se doporučit kterékoli z nich. Normální rozdělení

vykazuje velkou míru vyhlazení (*smoothing*), jak už z charakteru jádrové funkce vyplývá. Úzké pásmo jádrové funkce vede obecně k nízkému stupni vyhlazení (*under-smoothing*); spektrální mapa takové konfigurace působí podexponovaně. Širší pásma vytváří vyhlazení příliš velké (*over-smoothing*); spektrální mapa je přexponovaná. Jako ideální se jeví šířky pásma 0,05 a 0,10. Užší pásma jsou vhodná pouze pro jádrové funkce typu *GAUSSIAN* (normální převážení).

Eliminací okrajových jevů, docílenou aplikací *post-processing* filtru na integrovaný jádrový odhad, značně klesne průměrná integrovaná chyba. Zdaleka největšího efektu snížení odchylky roviny minimálního kontrastu se takto dočkáme u normálního převážení, kde už samotná spektrální mapa s ošetřenými okraji téměř splývá s vizualizací roviny cílové funkce délkové intenzity. Máme-li možnost korekci pomocí koeficientů aktivovat, je nasnadě tak učinit.

S rostoucím exponentem minimalizace kontrastu pozvolna roste i průměrná chyba. V reálných aplikacích si tedy můžeme bez obav vystačit s hodnotou exponentu rovnou jedné (minimalizace absolutní odchylky), případně volíme stupeň dva (metoda nejmenších čtverců). Nastavování vyšších exponentů již smysl prakticky nemá – odchylka roviny minimálního kontrastu narůstá a vlastní výpočet hledání minima je složitější.

Samotným závěrem uvedme, že předkládaná práce uspokojivě naplnila všechny stanovené cíle – vhodným nastavením vstupních parametrů jsme schopni dostatečně přesně aproximovat funkci intenzity, a to rovinou s minimálním kontrastem vůči jádrovému odhadu. Nabíledni je zamyšlení nad možností převedení výsledků simulačních experimentů do ryze praktické sféry; ve formě doplňku videometrie nanovlákněných struktur. Ze zcela odlišného oboru zoologického uvedme alespoň sledování pohybu divokých zvířat, zejména migraci ptactva. Další možností rozvoje tématu by mohla být tvorba obdobného simulačního prostředí pro složitější struktury prostorové statistiky; kupříkladu stochasticky generované trojúhelníky, čtyřúhelníky, obecně náhodné mnohoúhelníky. Vyšetřování délkové intenzity rovinných útvarů, na něž je možno aplikovat jádrový odhad, lze obecně poměrně snadno implementovat metodou minimálního kontrastu.

Použitá literatura

- [1] BENES, Viktor a Jan RATAJ. *Stochastic Geometry: Selected Topics*. Boston: Kluwer Academic Publishers, 2004, 221 p. ISBN 14-020-8103-0.
- [2] DIGGLE, Peter. A Kernel Method for Smoothing Point Process Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* [online]. 1985, vol. 34, no. 2, s. 138–147, 2011-10-03 [cit. 2013-05-01]. Dostupné z: <http://www.jstor.org/stable/2347366>
- [3] DUONG, Tarn. An introduction to kernel density estimation. In: DUONG, Tarn. *An introduction to kernel density estimation* [online]. 2001 [cit. 2013-05-01]. Dostupné z: <http://www.mvstat.net/tduong/research/seminars/seminar-2001-05/>
- [4] DUONG, Tarn. Plug-in Bandwidth Selectors for Bivariate Kernel Density Estimation. In: DUONG, Tarn. *Plug-in Bandwidth Selectors for Bivariate Kernel Density Estimation* [online]. 2002-05-30 [cit. 2013-05-08]. Dostupné z: <http://www.mvstat.net/tduong/research/seminars/seminar-2002-05-30.pdf>
- [5] DUPAČ, Václav a Marie HUŠKOVÁ. *Pravděpodobnost a matematická statistika*. Praha: Karolinum, 2003. ISBN 80-246-0009-9.
- [6] *ELMARCO: Nano for Life* [online]. © 2004–2013 [cit. 2013-05-08]. Dostupné z: <http://www.elmarco.cz/>
- [7] EPANECHNIKOV, V. A. Non-Parametric Estimation of a Multivariate Probability Density. *Theory of Probability* [online]. 1969, vol. 14, issue 1, s. 153-158 [cit. 2014-05-08]. DOI: 10.1137/1114019. Dostupné z: <http://epubs.siam.org/doi/abs/10.1137/1114019>
- [8] JIRSÁK, Čeněk. *Nestacionární procesy částic*. Praha, 2010. Diplomová práce. Univerzita Karlova v Praze, Matematicko-fyzikální fakulta, Katedra pravděpodobnosti a matematické statistiky. Vedoucí práce Doc. RNDr. Jan Rataj, CSc.
- [9] PAWLAS, Zbyněk. *Prostorové modelování, prostorová statistika I (NSTP005)* [online]. Praha, 2012, 2012-01-12 [cit. 2013-05-01]. Dostupné z: <http://www.karlin.mff.cuni.cz/~pawlas/2012/STP005/prostor.pdf>. Pracovní text k přednášce. Univerzita Karlova v Praze, Matematicko-fyzikální fakulta, Katedra pravděpodobnosti a matematické statistiky.
- [10] RANCOITA, Paola M.V., Alessandro GIUSTI a Alessandra MICHELETTI. INTENSITY ESTIMATION OF STATIONARY FIBRE PROCESSES FROM DIGITAL IMAGES WITH A LEARNED DETECTOR. *Image Analysis* [online]. 2011-11-18, vol. 30, issue 3, s. 167-178 [cit. 2014-05-28]. DOI: 10.5566/ias.v30.p167-178. Dostupné z: <http://www.ias-iss.org/ojs/IAS/article/view/908>
- [11] ROSS, Sheldon M. *Introduction to Probability Models: Tenth Edition*. 10. vyd. Boston: Academic Press, 2010, 784 p. ISBN 01-237-5686-3.

Přílohy

A Ukázka simulačního prostředí

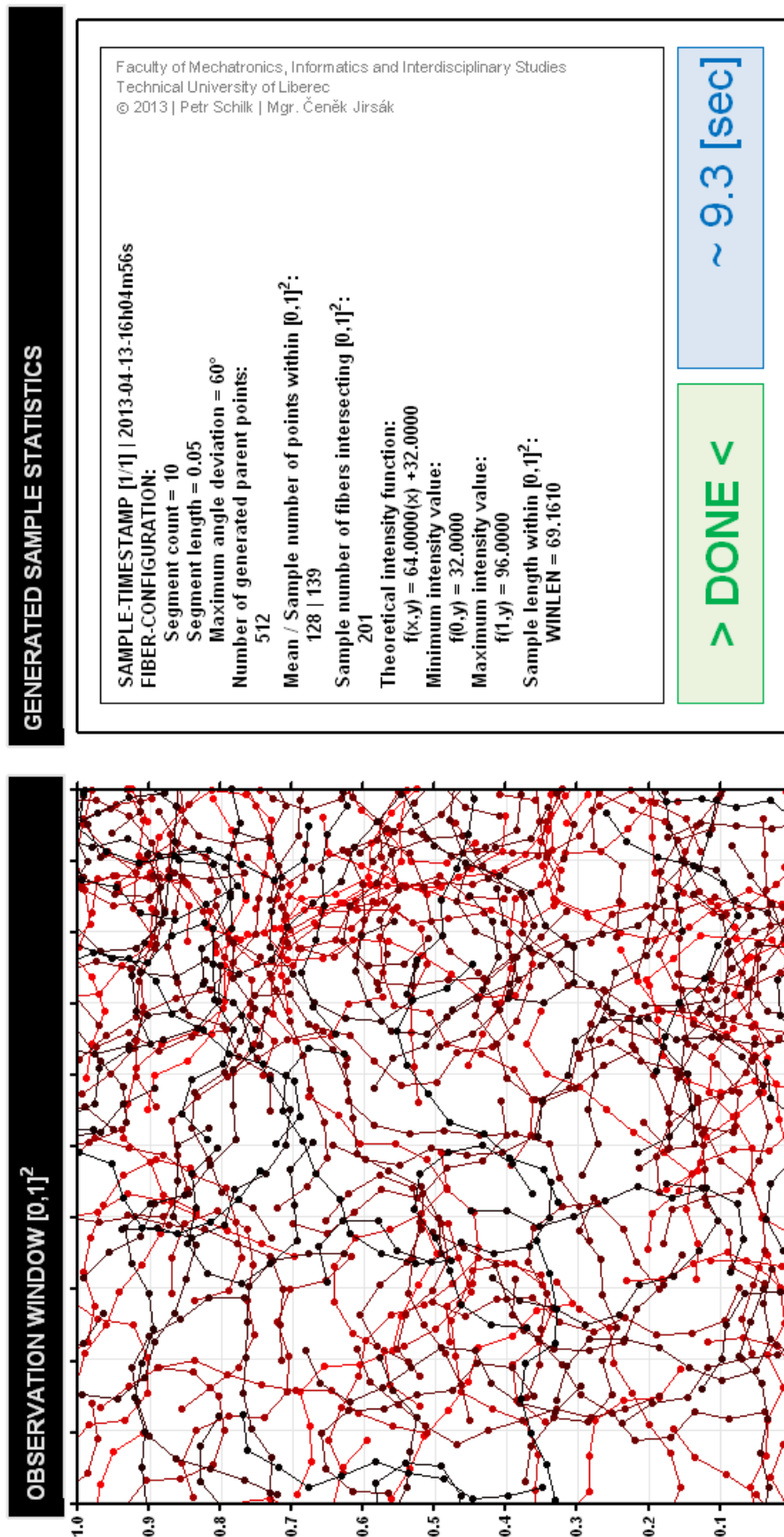
A.1 Parametry generovaného vzorku

```
=====
SAMPLE-TIMESTAMP | 2013-04-13-16h04m56s
=====
FIBER-CONFIGURATION:
> Segment count | 10
> Segment length | 0.05
> Maximum angle deviation | 60°
Number of generated parent points | 512
Mean / Real number of points within [0,1]^2 | 128 | 139
Sample number of fibers intersecting [0,1]^2 | 201
Theoretical intensity function |  $f(x,y) = 64.0000(x) + 32.0000$ 
> Minimum intensity value |  $f(0,y) = 32.0000$ 
> Maximum intensity value |  $f(1,y) = 96.0000$ 
Sample length within [0,1]^2 | WINLEN = 69.1610
Elapsed time | 9.3 [sec]
=====
```

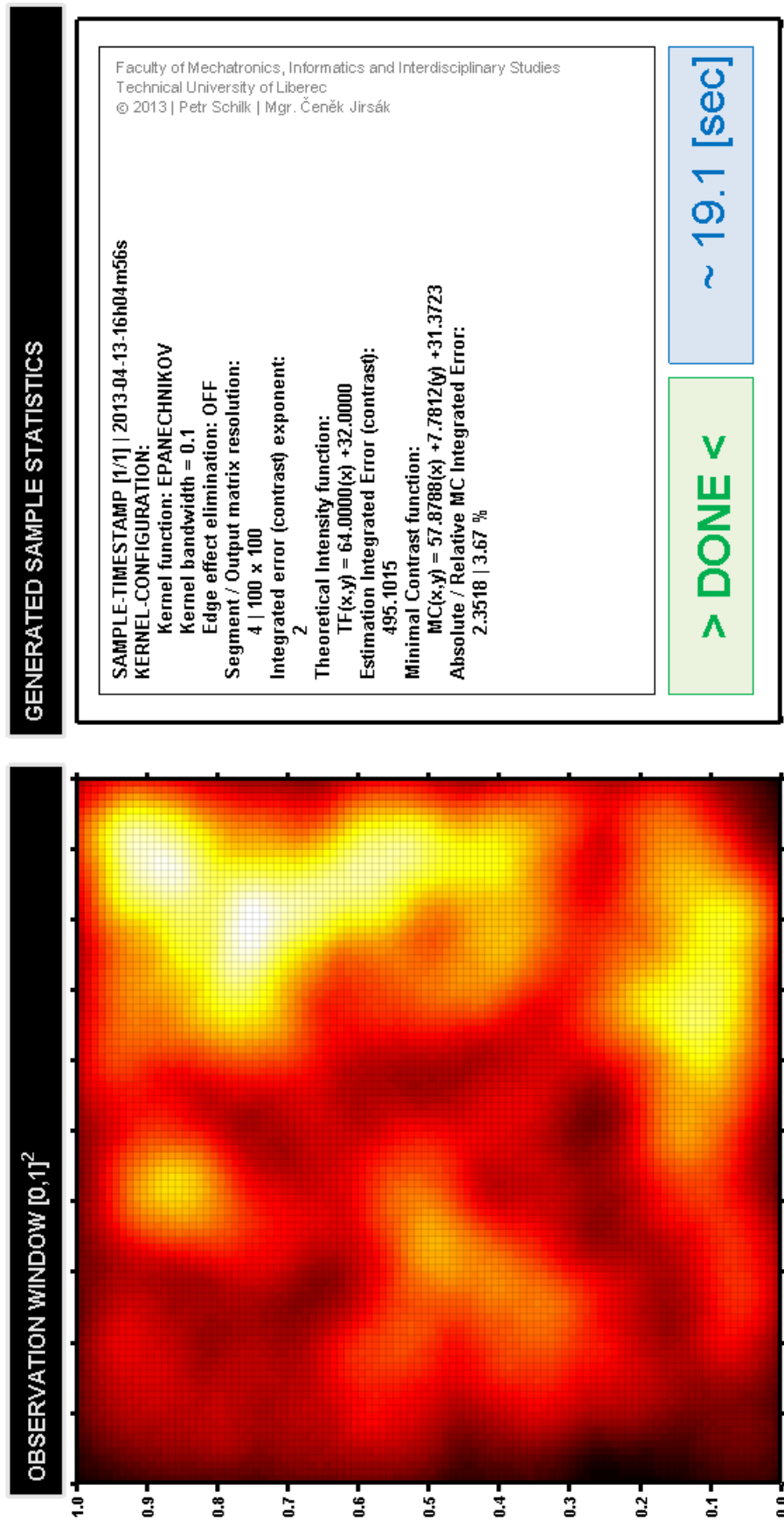
A.2 Parametry jádrového odhadu

```
=====
SAMPLE-TIMESTAMP | 2013-04-13-16h04m56s
=====
KERNEL-CONFIGURATION:
> Kernel function | EPANECHNIKOV
> Kernel bandwidth | 0.1
> Edge effect elimination | OFF
Segment / Output matrix resolution | 4 | 100 x 100
Integrated error (contrast) exponent | 2
Theoretical Intensity function |  $TF(x,y) = 64.0000(x) + 32.0000$ 
Estimation Integrated Error (contrast) | 495.1015 | 758.35 %
Minimal Contrast function |  $MC(x,y) = 57.8788(x) + 7.7812(y) + 31.3723$ 
Absolute / Relative MC Integrated Error | 2.3518 | 3.67 %
[null]
Elapsed time | 19.1 [sec]
=====
```

A.3 Generování vzorku procesu vláken



A.4 Aplikace jádrového odhadu



B Obsah přiloženého DVD

Přiložený disk DVD-ROM je členěn do následujících adresářů (složek):

- **01-BP_TEXT [5,18 MB]**
 - *2013-BP_Schilk Petr.pdf*
 - *2013-BP_Zadání.pdf*
 - *2013-BP_Prohlášení.docx*
- **02-BP_MATLAB [54,99 kB]**
 - naprogramované simulační prostředí
 - zdrojové kódy skriptů MATLAB:
 - scr1_fiberproc_gen.m*
 - scr2_kernel_est.m*
 - funkčnost testována v následujících verzích MATLAB:
 - 7.5.0.342 (R2007b) [2007-08-15]
 - 7.9.0.529 (R2009b) [2009-08-12]
- **03-BP_SIMULACE [1,54 GB]**
 - podrobné výsledky veškerých proběhnuvších simulačních experimentů
 - generované vzorky procesu vláken
 - spektrální mapy jádrových odhadů
 - výstupní textové soubory s konfiguracemi vzorků procesu vláken
 - výstupní textové soubory s vypočtenými ukazateli jádrových odhadů